

A SYNCHRONIZATION TECHNIQUE TO PROVIDE QUALITY OF SERVICE IN BROADBAND ISDN

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Engineering
in
Electrical and Electronic Engineering
at the
University of Canterbury
Christchurch, New Zealand
September 1995

Don Kalana Manohara Weerasiri
Bsc (Eng.) Hons

ABSTRACT

This thesis analyzes the different Quality of Service (QoS) parameters that are important in setting up multimedia calls in Broadband ISDN (B-ISDN). Unlike in traditional telecommunication networks, service providers of Broadband ISDN provide guaranteed Quality of Service. This requires a comprehensive QoS management framework. The users of B-ISDN are able to adapt the Quality of Service of multimedia to their requirements. The main adaptable QoS parameters include video parameters, audio parameters and synchronization parameters. Intra-media synchronization and inter-media synchronization are important aspects of overall QoS provision.

The audio and video streams of a particular source can be transmitted in different paths as these streams may have varying QoS requirements. Thus the audio and video streams will incur different transfer delays. The encoder and decoder processing delays of audio and video are also different. When presenting audio and video, the video stream should not be allowed to lag or lead the audio stream beyond acceptable skew tolerances. In specialized applications, multiple audio and video streams have to be played out in synchrony at the receiver. Also the layers of a layered coding system may be transmitted in different paths and the layers should be perfectly synchronized at the receiver.

Management of special resources such as layered encoders and decoders, media combiners code converters etc., ensures that intra-media and inter-media synchronization is maintained in isochronous communications. The different special resources and their relevance to QoS parameters such as intra-media and inter-media synchronization is discussed in detail.

In this thesis, a comprehensive synchronization mechanism called Static Delay Compensation (SDC) is proposed to guarantee synchronization QoS. When synchronizing two or more media streams static delays are added to streams with lower end-to-end delays. It will be shown that SDC will satisfy almost all of the synchronization QoS requirements.

ACKNOWLEDGMENTS

First and foremost I wish to express my sincere gratitude to my supervisor Dr. Chris Carey Smith for all the guidance and encouragement he has given me. His constructive criticism was very helpful in moulding this thesis. I am also thankful to him for providing me with financial assistance through Telecom New Zealand Limited.

I wish to thank the staff members and my postgraduate colleagues of the Department of Electrical and Electronic Engineering. Special thanks to Grant Naish for providing me with literature relevant to my research and for sharing his work. Static Delay Compensation was his initial idea. I thank Kelvin Morton for all the help given on Visual Prolog and Kevin Taylor for all the valuable discussions.

I wish to thank Prof. Harsha Sirisena and Dr. Felix Hartanto for giving their suggestions regarding this thesis. I am also thankful to them for giving me an opportunity to present my work at the network research group seminars.

Finally I thank the members of my family for all the help they have given me. I am grateful to my wife for all the support she has rendered. We had our first child during the period of research and we had to get adjusted to a new life. I wish to thank my parents and my wife's parents for all the encouragement and support they have given me in successfully completing this thesis.

CONTENTS

ABSTRACT		i
ACKNOWLEDGMENTS		iii
GLOSSARY		ix
Chapter 1	INTRODUCTION	1
	1.1 Broadband ISDN	1
	1.2 Quality of Service in Broadband ISDN	2
	1.3 The Contributions of this Thesis	3
Chapter 2	QUALITY OF SERVICE	5
	2.1 Definition	5
	2.2 Telecommunications QoS on Different Timelines	6
	2.2.1 QoS in Call Timeline	6
	2.2.1.1 QoS in Call and Connection Establishment	6
	2.2.1.2 QoS in Information Transfer	7
	2.2.1.3 QoS in Call and Connection Termination	8
	2.2.2 QoS in Contract Timeline	8
	2.2.3 QoS in Service Infrastructure Timeline	8
	2.3 Quality of Service Management	9
	2.3.1 QoS Translation and Negotiation	10
	2.3.2 QoS Verification	11
	2.3.3 QoS Maintenance	12
	2.3.4 Resource Management	12
	2.3.5 Signalling	13
	2.4 Layered Model for QoS	13
	2.5 Audio and Video Quality of Service	15

2.5.1 Audio Quality of Service	16
2.5.2 Video Quality of Service	16
2.5.3 Synchronization Quality of Service of Audio and Video	16
Chapter 3 SPECIAL RESOURCES	19
3.1 Special Resource Management	19
3.2 Resource Allocation	19
3.2.1 Simultaneous Allocation of Resources	20
3.2.2 Post Allocation of Resources	20
3.3 Code Converters	21
3.4 Multicasting Switches	22
3.5 Encoders and Decoders	22
3.5.1 Audio Encoders and Decoders	22
3.5.2 Video Encoders and Decoders	23
3.5.2.1 H.261	23
3.5.2.2 MPEG	24
3.6 Media Combiners	26
3.6.1 Combining Audio	26
3.6.2 Combining Video	27
3.6.2.1 Pel Domain Approach	27
3.6.2.2 DCT Domain Approach	28
3.6.2.3 VLC Domain Approach	29
3.6.3 Synchronization of Combined Streams	29
3.7 Layered Encoders and Decoders	30
3.7.1 Spatial Band Decomposition	31
3.7.2 Temporal Band Decomposition	32
3.7.3 Spatio-temporal Decomposition	33
3.7.4 Layered Coding Based on Human Visual Characteristics	33
3.7.5 Synchronization of Layered Coding Systems	34
3.8 Media Bridges	35
3.9 Summary	35

Chapter 4	SYNCHRONIZATION	37
	4.1 Sender and Receiver Clocks	37
	4.1.1 Clock Recovery Methods	38
	4.1.1.1 Synchronous Residual Time Stamp	38
	4.1.1.2 Adaptive Clock Recovery Method	38
	4.1.2 Slip and Pause	39
	4.2 End-to-end Delay of a Media Stream	40
	4.3 Intra-media Synchronization	40
	4.4 Examples of Total Jitter in End-to-end Delay	42
	4.5 Effect of Jitter on Slip/pause Mechanism	43
	4.5.1 Slip	43
	4.5.2 Pause	44
	4.6 Slip/pause Frequency	46
	4.7 Inter-media Synchronization	47
	4.8 Static Delay Compensation	49
 Chapter 5	 STATIC DELAY COMPENSATION ; A MECHANISM FOR SYNCHRONIZING TWO MEDIA STREAMS	 51
	5.1 Static delay Compensation (SDC)	51
	5.2 Case I : Both Skew Constraints Satisfied	52
	5.3 Case II : One Skew Constraint Satisfied and the Streams are able to be Synchronized	54
	5.4 Case III : One Skew Constraint Satisfied and the Streams are not able to be Synchronized	55
	5.5 Case IV : Both Skew Constraints not Satisfied	56
	5.6 Summary of All Possible Situations	57
	5.7 Examples of Application of Static Delay Compensation	58
	5.7.1 Example 1	58
	5.7.2 Example 2	59
	5.7.3 Example 3	60

Chapter 6	STATIC DELAY COMPENSATION ; A MECHANISM FOR SYNCHRONIZING THREE MEDIA STREAMS	63
	6.1 Synchronization of Three or more Media Streams	63
	6.2 Constraints to be Satisfied when Synchronizing Three Media Streams	64
	6.3 Derivation of Static Delay Value for Three Media Streams	70
	6.3.1 Case 1	72
	6.3.2 Case 2	73
	6.3.3 Case 3	75
	6.3.4 Case 4	76
	6.3.5 Case 5	77
	6.3.6 Case 6	79
	6.3.7 Summary of All Possible Situations	80
Chapter 7	STATIC DELAY COMPENSATION ; A MECHANISM FOR SYNCHRONIZING MULTIPLE MEDIA STREAMS	83
	7.1 Constraints to be Satisfied when Synchronizing n Media Streams	83
	7.2 Static Delay Value for n Media Streams	84
	7.3 Example of Synchronizing Multiple Audio and Video Streams	87
	7.4 Synchronization of Layered Coding Systems	89
	7.5 Example of Synchronizing a Layered Coding System	91
Chapter 8	DISCUSSION AND CONCLUSION	93
	8.1 Discussion	93
	8.2 Conclusion	95
	8.3 Future Considerations	96
	REFERENCES	97

GLOSSARY

ADPCM	Adaptive Differential Pulse Code Modulation
ATM	Asynchronous Transfer Mode
BC	Bearer Control
BWM	Bandwidth Management
B-ISDN	Broadband Integrated Services Digital Network
CC	Call Control
CCITT	Consultative Committee of International Telegraph and Telephone
CIF	Common Intermediate Format
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Code Modulation
DSI	Digital Speech Interpolation
FDM	Frequency Division Multiplexing
GOB	Group Of Blocks
GN	Group Number
HP	High Priority
HVS	Human Visual System
JPEG	Joint Photographic Experts Group
LSB	Least Significant Bits
LP	Low Priority
MCU	Multipoint Control Unit
MPEG	Moving Picture Experts Group
MSB	Most Significant Bits
MTBF	Mean Time Between Failure
MTTR	Mean Time To Repair
N-ISDN	Narrowband Integrated Services Digital Network
PCM	Pulse Code Modulation
POTS	Plain Old Telephone Service
PSTN	Public Switched Telephone Network
QCIF	Quarter Common Intermediate Format
QMF	Quadrature Mirror Filter
QoS	Quality of Service

RC	Resource Control
SKF	Short Kernel Filter
SRM	Special Resource Management
SRTS	Synchronous Residual Time Stamp
TDM	Time Division Multiplexing
VBR	Variable Bit Rate
VLC	Variable Length Coding

CHAPTER 1

INTRODUCTION

This chapter introduces the Broadband Integrated Services Digital Network (B-ISDN) and then presents the characteristics of Quality of Service (QoS) in B-ISDN. It then outlines the contributions of this thesis.

1.1 Broadband ISDN (B-ISDN)

Today's networks can be characterized by their specialization [De Prycker, 1993]. The telex network provides for 300 bits/s character transmission. The Plain Old Telephone Service (POTS) which is facilitated via the Public Switched Telephone Network (PSTN), provides for two way 4 KHz voice transmission. Television broadcast is transmitted through VHF / UHF frequencies or coaxial cable. In the private domain, data is transmitted through Local Area Networks (LAN). A network designed for a specific service can not be efficiently adopted by another, due to high specialization of these networks. A classic example is transmission of computer data across PSTN based on the X.25 protocol, where many Digital to Analog (D/A) and Analog to Digital (A/D) converter stages are used.

A first step towards a single universal network, was the introduction of N-ISDN, in which voice and data are transported over a single medium. This standard eliminated the inefficiency created by the Digital to Analog (D/A) and Analog to Digital (A/D) converter stages in data transmission. Digital N-ISDN switches are designed for 64 Kbits/s circuit switched channels [Handel, 1992]. This offers only a limited bandwidth granularity i.e. even if a bearer connection requires much less than 64 Kbits/s still a circuit of capacity 64 Kbits/s has to be reserved.

When designing B-ISDN one must take into account existing as well as future services. Bandwidth granularity should be provided for unknown services as well. The primary design criteria in B-ISDN was to make it the most flexible network in terms of bandwidth

allocation and the most efficient in terms of resource usage. Packet switching technology was adopted for B-ISDN as it readily provided for above mentioned attributes.

X.25 was a packet switching network designed for not so reliable underlying transmission links. Frame relay was designed for packet switching across N-ISDN, and it had less protocol processing at the switches. However, neither X.25 nor frame relay are suitable for isochronous services because of the complexity of the switching nodes. Therefore a new packet switching transfer mode called Asynchronous Transfer Mode (ATM) was introduced for B-ISDN. ATM only needs limited protocol processing functionality at switching nodes as the underlying physical transmission medium is reliable. ATM facilitates transmission of isochronous services due to its low switching delay.

1.2 Quality of Service in Broadband ISDN

Unlike traditional telecommunication networks, B-ISDN will provide adaptable and guaranteed QoS. QoS performance guarantee is facilitated by the utilization of fibre optic media as the physical transfer medium and ATM as the underlying transfer mode. The jitter in the propagation delay across fibre optic networks is almost negligible while ATM switches introduce only a small delay-jitter of about 250 μ s per switch [Bauer *et. al.*, 1995]. The above performance guarantees, at both the physical layer as well as at the ATM layer facilitate a very high QoS.

The users of B-ISDN are able to negotiate for the QoS which is to be provided. This includes negotiating both the QoS of the man-machine interface as well as the end-to-end QoS [Iai *et. al.*, 1993]. The QoS of the man-machine interface includes the video quality parameters such as video frame resolution, aspect ratio, frame rate, number of colours and also audio quality parameters such as mono/stereo reception, hi-fidelity/ telephone quality reception. End-to-end QoS parameters include end-to-end delay, jitter in the end-to-end delay, cell loss rate of the media stream, intra-media and inter-media synchronization between media streams.

Guaranteeing the QoS provision requires a comprehensive QoS management framework. In B-ISDN a multitude of special resources such as media combiners or multicasting switches must be managed, whereas in traditional telecommunication networks the only resource that had to be managed was the bandwidth resource. A special resource management entity is a very important aspect of Quality of Service management.

The bearer connections for which the Quality of Service has been negotiated have to be set up by the service provider. Audio and video streams of a particular source or the layers of a layered coding system may be transmitted in different bearer connections as they have different QoS requirements. If media with different QoS are transmitted across a single bearer connection the overall QoS of the bearer connection has to be that pertaining to the highest QoS. This results in a sub-optimal allocation of resources.

Synchronization is an important characteristic of overall QoS provision. The audio and video streams as well as layers of a layered coding system have to be played out in synchrony at the receiver. The performance guarantees of both the fibre optic physical layer and the ATM layer provides a solid foundation for guaranteeing the synchronization Quality of Service at the presentation layer.

1.3 The Contributions of this Thesis

This thesis investigates the Quality of Service parameters that are negotiable and adaptable. It also investigates the requirements that should be satisfied in guaranteeing the provision of Quality of Service. Chapters 2 and 3 summarize background information in the field of research and chapters 4 to 7 contain the author's original work.

In chapter 2, the concept of Quality of Service is discussed in detail. Initially the philosophical point of view of this concept is outlined and then the QoS parameters that are applicable for different timelines is described. A Quality of Service management framework that could be utilized to provide guaranteed QoS is suggested. The QoS provided by different audio and video standards is presented. The chapter concludes with a description of synchronization Quality of Service.

A detailed description of special resources that may need to be configured when setting up a call in Broadband ISDN is presented in chapter 3. A special resource is exclusively reserved for a call at a given time, unlike the total bandwidth which has to be shared between a number of calls. Special resource management mainly deals with timing and synchronization constraints of media streams.

Chapter 4 discusses the topic of synchronization. The problem of sender and receiver clock asynchrony in packet switched networks is discussed initially. The different delays that

contribute to the end-to-end delay of a media stream are analyzed and a mechanism is proposed to achieve intra-media synchronization in the presence of both the delay-jitter introduced by the network and the asynchrony of the encoder and decoder real-time clocks. Chapter 4 also introduces the concept of Static Delay Compensation (SDC) which is a new technique that could be utilized in facilitating inter-media synchronization of media streams. SDC is compared with other inter-media synchronization protocols to determine its merits in carrying out inter-media synchronization.

A method to achieve inter-media synchronization of two media streams by the application of the Static Delay Compensation technique is proposed in chapter 5. Few cases need to be considered when synchronizing two media streams. The streams may be already synchronized if the maximum lead and maximum lag of one stream with respect to the other is within the corresponding skew tolerance values. If the streams are not already synchronized they may be synchronized by the addition of a static delay to the stream with lower end-to-end delay. There may be situations where the streams cannot be synchronized.

In chapter 6, synchronization between three media streams is investigated. In synchronizing three media streams, every combination of two streams has to satisfy the skew constraints. In addition to this, every combination of three streams has to satisfy other constraints. In this chapter, a generalised equation to calculate the optimum static delay of a stream when synchronizing three media streams is derived.

In chapter 7, the method used to synchronize three media streams is generalized, firstly to derive the constraints that are applicable when synchronizing n media streams and secondly to calculate the static delay of any stream when synchronizing n media streams. The use of Static Delay Compensation to provide synchronization in layered coding systems is also described in chapter 7. Chapter 8 presents the conclusion and further research that should be carried out on the area covered in this thesis.

CHAPTER 2

QUALITY OF SERVICE (QoS)

This chapter gives a detailed description of the concept of Quality of Service (QoS). Quality as perceived by mankind is subjective and any treatment on QoS should outline the philosophical aspects associated with it. A comprehensive QoS management framework is necessary in guaranteeing the provision of QoS. The QoS management incorporates many sub-processes. A structured layered model is adopted for QoS management.

2.1 Definition

Quality of Service could be defined as the following. “Quality of Service (QoS) is a set of user perceived attributes that makes a service what it is. It is expressed in user understandable language and manifests itself as a number of parameters, all of which have either subjective or objective values.” [RACE QOSMIC deliverable D 1.3C, 1992]. As the definition suggests, the quality of a product is a philosophical issue and the following discussion highlights the intricacies involved in defining quality. The absolute or the abstract quality of a product could only be perceived by the divine and any person could only perceive a projection of the absolute quality of a product. Thus, quality depends on the person who experiences it.

Mr. A. may perceive a particular subset of quality parameters of the product and Mrs. B. may perceive a different subset of quality parameters of the same product. Therefore Mr. A’s quality of a product may differ from Mrs. B. Furthermore if Mr. A. and Mrs. B. perceive the same quality attribute in the same manner their evaluations of the quality may differ. Therefore Mr. A. and Mrs. B. may have different quality interpretations of the same product and the quality as perceived by humans is subjective [Mourelatou, *et. al.*, 1994]. Objective quality parameters are also a subset of absolute quality and are a set of parameters defined to help two or more human beings to perceive that same quality attribute in the same manner.

In any engineering discipline the subjective quality parameters are reduced to objective quality parameters in order to specify them. Specification of quality parameters helps to evaluate, verify and maintain them. Quality of Service (QoS) refers to such objective quality parameters pertaining to a telecommunication service.

2.2 Telecommunications QoS on Different Timelines

In telecommunications, the QoS attributes are specified in three different chronological dimensions; the call timeline, the contract timeline and the service infrastructure timeline. [RACE QOSMIC *deliverable D 1.3C*, 1992].

2.2.1 QoS in Call Timeline

The call timeline consists of the phases of

- call and connection establishment
- information transfer
- call and connection termination

of a telecommunication call.

2.2.1.1 QoS in Call and Connection Establishment

Call blocking probability during call establishment directly reflects the quality of service provided to the customer. Call establishment will incur a delay from the time the telecommunication user dials in, to the time the call is connected. A lower call establishment delay contributes to a higher QoS. If a customer experiences a delay variation when establishing a call to the same destination over a few call attempts that would also signify a lower QoS.

In B-ISDN a multimedia call is made up of different bearer connections pertaining to video, audio and data. A user may want to add a bearer connection or modify the quality of a bearer connection after the call is established. Also the network may need to downgrade the quality of a bearer connection. Thus a requirement exists for bearer connection establishment or

bearer connection re-establishment, separate from call management. Any connection re-establishment delay and delay variation also reflect a reduced QoS provision to the user.

If the cost of establishing a call with particular characteristics is low, it would signify a higher QoS. In this situation, although the QoS of information transfer is the same, the call can be established through different physical links and switches by different service providers. As a result, some service providers may be able to charge lower rates.

2.2.1.2 QoS in Information Transfer

QoS in information transfer can be divided into adaptable and non-adaptable QoS parameters. The non-adaptable QoS parameters are the parameters which cannot be controlled by the user. Premature termination probability, information loss and error rate of bearer connections are non-adaptable QoS parameters. The non-adaptable QoS parameters are similar to QoS parameters that existed in traditional networks.

With B-ISDN the users are able to adapt the Quality of Service according to their own requirements and this facility may be provided for all different bearer connections. The adaptable QoS parameters during the information transfer phase can be divided into two categories. They are QoS of the man-machine interface and the end-to-end QoS.[Iai *et. al.*, 1993]

The QoS of the man-machine interface can be further divided into audio, video and synchronization quality. The video QoS includes the horizontal and vertical frame resolution, aspect ratio of the video frame, the video frame rate, number of colours, colour / black & white reception, and number of layers to be decoded in a layered coding system. The audio QoS includes the bandwidth of the audio signal, the bit rate of the audio stream and mono/stereo reception. The synchronization QoS parameters include intra-media synchronization of a media stream, inter-media synchronization of media streams such as lip synchronization of audio and video, and synchronization between layers of a layered coding system. The QoS parameters of the man-machine interface are fully adaptable within the limits of underlying service provision.

The end-to-end QoS includes the allowable end-to-end delay and jitter of a media stream, and cell loss of the media stream. A lower end-to-end delay signifies a higher QoS in interactive communications. The speed of data transfer is also an end-to-end QoS parameter.

2.2.1.3 QoS in Call and Connection Termination

In call or connection establishment, resources are allocated to serve the call or connection and the resources remain allocated statistically or exclusively throughout the duration of the call. When call or connection termination is initiated, the service provider should deallocate the resources that were committed. If there is a delay in deallocating the resources the user may have to bear extra cost and also the network provider may incur higher call blocking probability. Thus call or connection termination delay is an important QoS parameter. If there is a delay variation during call or connection termination that would signify a lower QoS.

2.2.2 QoS in Contract Timeline

The contract timeline represents a long term contract between the customer and the service provider and consists of subscription, in-service and cancellation phases. The QoS parameters of this timeline can be categorized into system reliability parameters, efficiency and accuracy parameters and courtesy parameters. System reliability parameters include factors such as Mean Time Between Failure (MTBF) and Mean Time to Repair (MTTR). Speed of provisioning the service and billing accuracy belong to efficiency and accuracy parameters. Courtesy parameters include activities such as help desk functions.

2.2.3 QoS in Service Infrastructure Timeline

Service infrastructure timeline refers to the development of the telecommunication service and consists of installation, operation and discontinuation phases of the telecommunication network. The QoS parameters in the service infrastructure timeline includes factors such as the technology being used for the telecommunication service, upgrading and development of the telecommunication service, and viability of the telecommunication service provider.

2.3 Quality of Service Management

The user specifies certain QoS parameters for a call or a bearer connection and expects the service provider to guarantee these QoS parameters. The primary objective of QoS management is to guarantee the delivery of QoS to the user by means of a *service contract* [Campbell *et.al.*, 1994]. The service contract contains parameters such as audio, video and synchronization QoS of bearer connections (section 2.5), commitment on bearer connections and clauses to specify the action to be taken when QoS violations occur.

The QoS commitment on bearer connections can be deterministic, statistical or best-effort. Deterministic or guaranteed commitment is based on fixed resource allocation where no resource gain is feasible. Statistical commitment is based on a shared resource allocation which encourages a higher degree of resource utilization. ATM is an asynchronous time division multiplexing network which provides statistical QoS commitment. Guaranteed QoS commitment can be easily facilitated by techniques such as reservation of peak bandwidth for bearer connections.

In best-effort commitment, no resources are allocated and for this type of commitment, resource allocation is only possible after other levels have been serviced. If only a best-effort or a statistical QoS performance is required then it is unnecessary to define a structured QoS management framework, although the provision of statistical or best effort services is not inhibited in such a structured framework. Thus, in guaranteeing the Quality of Service, the QoS manager has to manage the heterogeneity of user requirements.

The QoS manager should also handle the portability of different platforms and dynamic resource changes. The different platforms the QoS manager should handle, include the variety of operating systems and special resources such as encoders, decoders and media combiners etc., which may belong to different standards. The availability of bandwidth resource would depend on the congestion level of the network and congestion level in turn would depend on the burstiness of the sources. Thus, the bandwidth is a dynamic resource, which needs to be managed by the QoS manager. The QoS manager should also ensure the satisfaction of time constraints of isochronous communications [Nahrstedt *et. al.*, 1995a].

The Quality of Service Management framework has to integrate a number of processes in order to achieve its objective. The figure 2.1 depicts these processes.

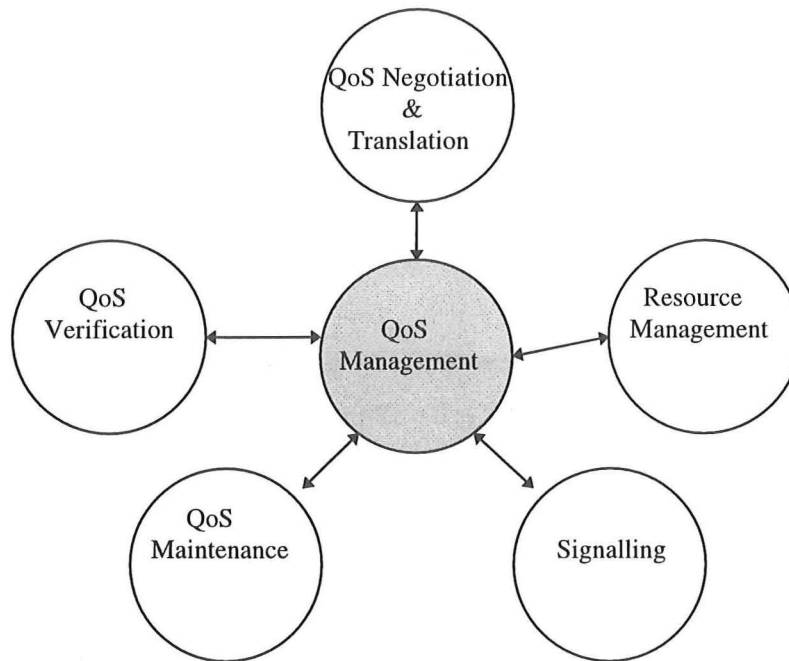


Figure 2.1 : QoS Management

2.3.1 QoS Negotiation & Translation

The purpose of QoS negotiation is twofold. Firstly it helps to establish common QoS parameter values among the service users and providers. Secondly it provides a means of capitalizing scarce resource capacities by reserving only the real demand at any point in time [Nahrstedt *et. al.*, 1995b]. The result of a QoS negotiation is a service contract between the user and the service provider. Once the service contract has been established it functions as a QoS specification and both the network and the user should adhere to it. The terms of the service contract is also called *user QoS*. The user QoS is not directly perceivable by the network and needs to be translated into network QoS [Jung *et. al.*, 1993] .

The QoS translation function translates the user QoS parameters into QoS performance targets at different layers [Hu, 1995]. Negotiations should be carried out to guarantee these performance targets by the architectural layers in association with resource management. For example, initially the user would inform his/her desire to establish a multimedia call with specified QoS parameters consisting of video, audio and synchronization characteristics. The network provider translates the above QoS parameters into performance

targets at the lower architectural layers. The possibility of guaranteeing the delivery of the performance targets has to be determined by assessing the capabilities of video and audio codecs and the transmission links.

QoS negotiations are carried out for *required values*, which may take the form of upper bound and lower bound values or mandatory values [Hutchison *et. al.*, 1994]. Figure 2.2 depicts a possible illustration of these values. A mandatory value is a level of QoS that must be maintained during a given percentage of time.

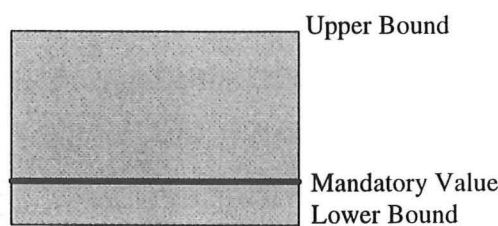


Figure 2.2 : Required values in QoS negotiation

The QoS negotiation and translation function must also provide for re-negotiation of quality of service which could be either network initiated or user initiated. If the user wants to modify the QoS after the information transfer begins, such as downgrading the QoS of audio from G.722 to G.721 then it would be a user-initiated re-negotiation. If the network cannot provide the negotiated QoS either due to a link failure or due to congestion and requires to modify the QoS of the call then it would be a network initiated re-negotiation. For example, the network may ask the user to downgrade the frame resolution from CIF to QCIF [CCITT H.261 recommendation]. In order to facilitate, network-initiated QoS re-negotiation, a QoS degradation path is also defined in the service contract. This specifies whether to ignore, inform the user or re-configure to the user specified level if a QoS violation occurs.

2.3.2 QoS Verification

The QoS verification function monitors the performance at various layers to see whether the performance targets are met [RACE QOSMIC *deliverable D 1.3C*, 1992]. The verification of user QoS is subjective. Subjective QoS verification has to be done through user surveys [Mourelatou, 1994]. However, due to the subjective nature of assessment, user surveys may not be the ideal method of QoS verification. Objective verification of QoS has to be carried

out by measuring network performance. In robust networks consisting of ATM and fibre optic links QoS parameters such as end-to-end delay and jitter are bounded, and performance of these parameters need not be verified. However, QoS parameters such as cell loss across transmission links may need to be verified. It is expensive to carry out evaluations of network performance for different calls in isolation, due to the need for specialized equipment and protocol processing.

2.3.3 QoS Maintenance

The functionality of QoS maintenance is twofold. Firstly it does policing which is used to ensure that the user does not violate the QoS by sending a higher information rate than what was negotiated. Secondly it evaluates causes of QoS violations in the network due to link failure or congestion and tries to rectify them. In a network where multiple calls are supported, it may not be easy to identify the cause for a QoS violation of a particular multimedia call, as explained in the previous section. Mourelatou *et. al.* [1994] discusses an object oriented approach to evaluate causes for QoS degradations. Each time a new bearer connection is set up, a dedicated object for monitoring the QoS of the bearer connection is created. This object keeps measurements and statistics of the performance parameters identified as important QoS aspects. This function may call for a QoS re-negotiation if the network is unable to deliver the negotiated QoS.

2.3.4 Resource Management

Resource Management is the process of managing the network resources as well as end-system resources. In telecommunication networks that evolved before Broadband ISDN, bandwidth was the only resource considered. Even in B-ISDN, the term “resource” is often referred to as bandwidth. However other resources such as layered encoders, layered decoders, code converters, multicasting switches, media combiners and media bridges also need to be considered in Broadband ISDN. In this research all the above are collectively referred to as “special resources” although in B-ISDN signalling proposals [RACE MAGIC deliverable 10 A, 1994 ; Moyer *et. al.*, 1993] this term does not include layered encoders and layered decoders. In addition to bandwidth and special resources, resource management should also manage computational resources such as CPU, main memory and I/O buffers.

A resource can be used exclusively by one process or shared among processes at a given time. A special resource is an exclusive resource whereas bandwidth is a shared resource. The resource management function [Nahrestdt *et. al.* , 1995b]

- Performs reservation and allocation of resources during multimedia call establishment, so that traffic can flow according the QoS specification.
- Must adhere to resource allocation during multimedia delivery.
- Adapts to resource changes during an ongoing multimedia session.
- Deallocates resources at multimedia call termination.
- Ensures intra-media and inter-media synchronization at special resources in isochronous communications.

2.3.5 Signalling

Signalling is the message transfer syntax used for QoS negotiation and translation as well as resource management between and across architectural layers. In both Plain Old Telephone Service and Narrowband ISDN the signalling protocol was monolithic. However, in the B-ISDN, functionalities of Call Control (CC), Resource Control (RC) and Bearer Control (BC) protocols are being defined [RACE MAGIC *deliverable 10 A*, 1994]. Unlike in previous telecommunication standards, B-ISDN has to configure different bearer connections pertaining to multimedia calls. This requires the separation of Bearer Control and Call Control. The Bearer Control functionality controls the allocation and deallocation of the bandwidth resources. Resource Control is defined to manage special resources independently from both Call Control and Bearer Control [Paglialunga *et. al.*, 1994].

2.4 Layered Model for QoS

Considering the diverse hierarchical functionalities involved, a layered model can be adopted for QoS Management. Each layer is represented by a *brick* as shown in figure 2.3 [RACE QOSMIC *deliverable D1.3C*] . Each brick negotiates QoS in association with resources pertaining to that layer. A brick has to satisfy the requirements of the upper layer as well as its peer. The brick generates the QoS requirements that should be satisfied by its

lower layer brick as well as its peer. Before the connection has been set up every low layer brick has to guarantee to the higher layer brick as well as to the peer brick that their requirements will be satisfied.

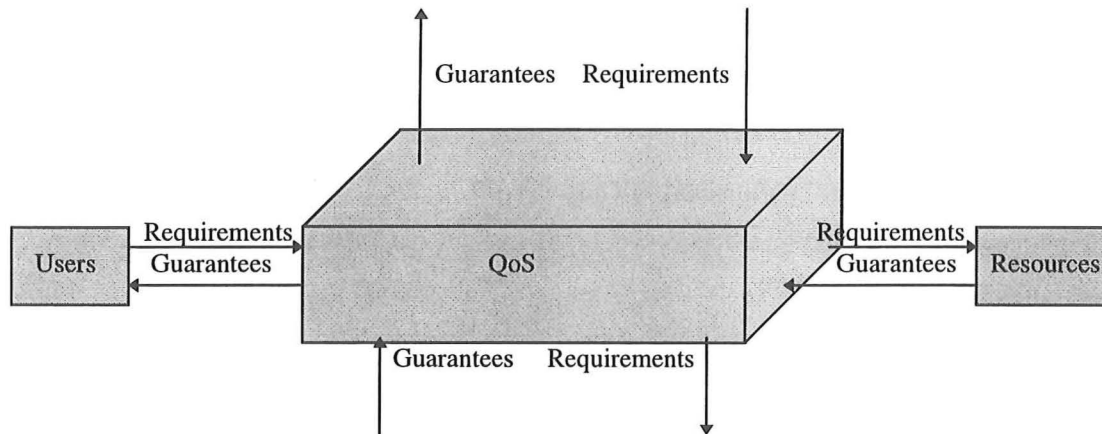


Figure 2.3 : QoS Brick

Figure 2.4 shows different layers that are applicable in the QoS management framework. The user QoS may be configured by a GUI-based tuning process in association with presentation devices such as monitors and microphones[Nahrstedt *et. al.*, 1995a]. The users could be presented with a test video stream to select the frame size, number of colors, frame rate etc. A test audio stream may be played to select the audio quality. Alternatively, QoS tuning may be done after the connection is established. If one wants a lower or raise QoS after the call has been established he/she could immediately call for a re-negotiation.

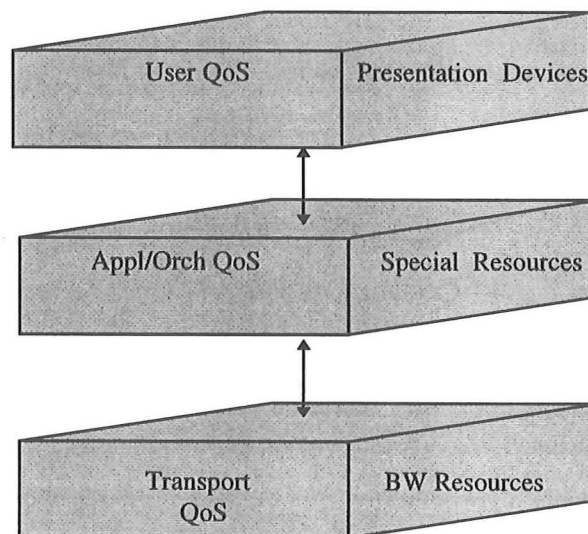


Figure 2.4 : Layered model for QoS

If two users are communicating using a telecommunication service it can be represented by a model as shown in figure 2.5. The signalling functionality at the user level is called Call Control (CC). The signalling functionality at the application layer/ orchestration layer and transport/network layers are called Resource Control (RC) and Bearer Control (BC), respectively. The application/orchestration layer configures QoS in association with Special Resource Management (SRM). The transport layer configures QoS in association with Bandwidth Resource Management (BWM).

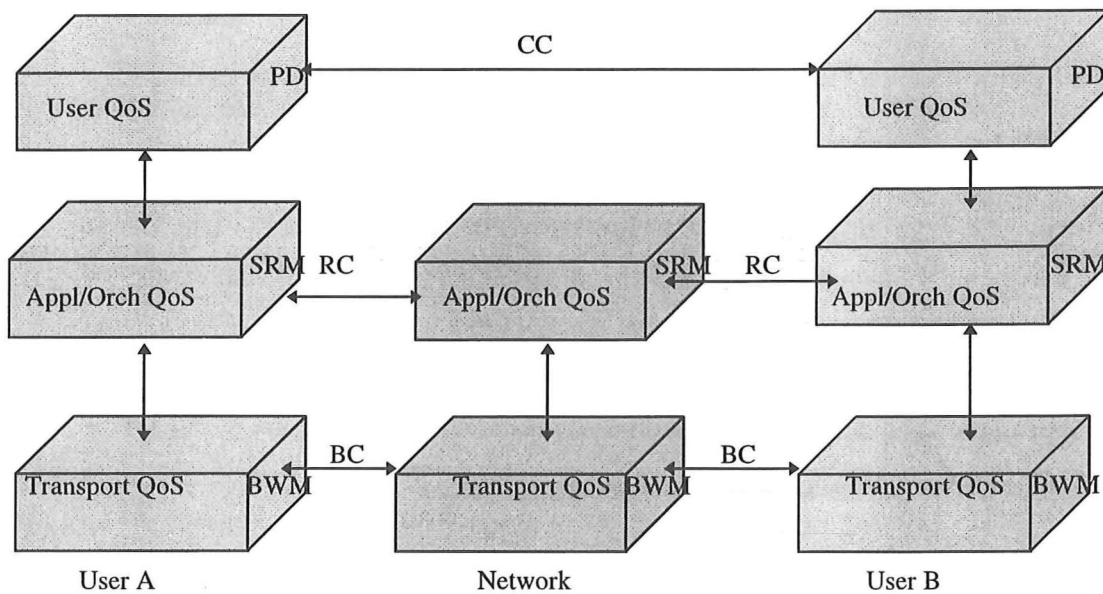


Figure 2.5 : Layered model for two communicating users

The application/orchestration layer as well as the transport layer functionalities are also present within the network. If a special resource such as a media combiner has to be located then the application/orchestration layer becomes active in the network [Mouchtaris *et. al.*, 1994]. The transport layer is always active within the network as the bandwidth resource is always being utilized.

2.5 Audio and Video Quality of Service

In this section, the QoS provided by different audio and video standards are outlined. The encoding and decoding processes of these standards are elaborated in section 3.5.

2.5.1 Audio Quality of Service

The end-to-end delay of an audio stream should be less than 400 ms for good interactive communications [Dimolistas *et. al.*, 1993 ; Zebarth, 1993]. The users could select telephone quality (0.3 - 3.4 KHz), wideband quality (0.05 -7 KHz) or hi-fidelity quality (0.02-20 KHz), for audio communication, depending on their hardware capabilities. G.711, G.721 and G.728 are the standards providing telephone quality communication and G.722 is a wideband audio quality standard. Hi-fidelity codecs are not standardized yet. The encoding and decoding processes of these standards are further elaborated in section 3.5.1.

2.5.2 Video Quality of Service

The users could select the quality of the video depending on the capabilities of the video codec being used for communication. H.261 and MPEG are the two main video codec standards used. The H.261 standard provides the flexibility of two video frame sizes, namely

- QCIF which has a frame size of 144 x 176 pixels
- CIF or FCIF which has a frame size of 288 x 352 pixels

The alternative name of H.261, px64 derives from a transmission rate formula of bandwidth equivalent to an integral multiple of 64 Kbps. If p equals to 1 or 2 the quality of reception would be poor to moderate and primarily this quality is appropriate for a videophone using QCIF images [Biggar *et. al.*, 1990]. If higher quality is required, $p \geq 6$ transmission is used. Moving Picture Experts Group (MPEG) standard facilitates transmission of high quality pictures with better frame resolutions. Section 3.5.2 elaborates the above two standards in detail.

2.5.3 Synchronization Quality of Service of Audio and Video

In playing multimedia at the receiver the video and audio should be lip synchronized. In traditional analog transmission methods, audio and video were multiplexed together and the transmission delay of both streams were identical. Therefore in analog television systems synchronized presentation of audio and video was an easy task.

Audio and video streams generated by a source may have different QoS requirements. If a particular stream requires a higher QoS than another stream originating from the same

source, then the streams may be transmitted across different paths. The stream with a higher QoS may be transmitted across a high QoS path while the other stream may be transmitted across a low QoS path. If the low QoS stream is also transmitted across the high QoS path, unnecessary allocation of the valuable resources of the high QoS path would occur. Thus if the audio and video streams have different QoS requirements it is always advantageous to transmit the two streams over different paths. This results in a different transfer delay of audio and video. Also the encoder and decoder processing delays of audio and video are different. Thus a mechanism should be in place to synchronize audio and video streams at the receiver.

The audio and video need not be played out in perfect synchrony. The playout of video can lead the playout of audio by a certain time and vice versa. This lead or lag is called skew tolerance. Humans are able to detect lack of synchronization of audio and video down to 30 ms. However the tolerable skew can be much larger [Graham, 1965]. In tests conducted by Cooper [1988] and Iai *et. al.*[1993], lead/lag thresholds of video over audio of 120 ms / 40 ms and 240 ms / 100 ms were recommended. In the subjective experiments carried out by Steinmetz *et. al.*[1993], the results shown in figure 2.6 were tabulated.

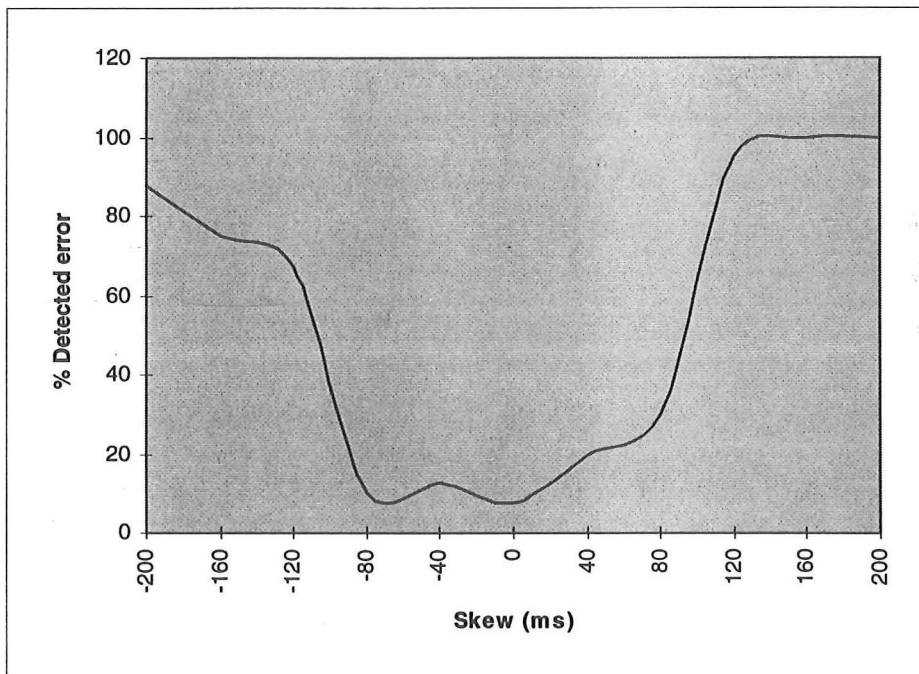


Figure 2.6 : Human visual perception of audio and video synchronization, the negative values represent a video lead while the positive values represent an audio lead

The human system is more accustomed to video leading audio, as light propagation is faster than acoustic wave propagation. If a person sits 20 m from the loudspeaker in a musical show, he/she will experience a video lead of 60 ms [Naish, 1994 b]. Thus, in setting up calls consisting of video presentations, the sitting distances of viewers should also be taken into consideration, in determining allowable skew through the network. According to the graph depicted in figure 2.6, a 90 ms video lead and a 60 ms audio lead produces a detected error below 20 %. The above figures will be used as conservative skew tolerance values for lip synchronization between audio and video, for a person sitting close to a display terminal.

CHAPTER 3

SPECIAL RESOURCES

Resource management is one of the main functionalities of QoS management. In Broadband ISDN, there are a number of special resources that should be managed such as layered encoders and decoders, media combiners, multicasting switches, code converters and media bridges. In this chapter, special resource management in B-ISDN is introduced and then a detailed description of special resources and their relevance to QoS parameters such as intra-media synchronization and inter-media synchronization is presented.

3.1 Special Resource Management

The functionality of special resource management in guaranteeing QoS is two fold. Firstly it deals with allocation and de-allocation of special resources. Secondly it ensures intra-media and inter-media synchronization of isochronous bearer connections. The special resource allocation may need to be changed if the QoS of the call or connection is modified. A special resource is an exclusive resource, as it will be allocated exclusively to a particular call at a given point in time. In the next section, the methods of resource allocation are described in detail. The special resources and their QoS characteristics are presented from section 3.3 onwards.

3.2 Resource Allocation

There are two possible ways in which resources could be allocated for a multimedia call in Broadband ISDN [RACE MAGIC deliverable 10A], namely simultaneous allocation and post allocation. In figure 3.1, layered protocol functionality between two communicating users are presented. In simultaneous allocation, resources are allocated concurrently with end-to-end call set up. Thus, Call Control (CC) functionality becomes concurrently active with Resource Control (RC) and Bearer Control (BC) functionalities. In post allocation,

end-to-end call establishment is carried out before resources are allocated, and RC and BC functionalities become active only after user QoS is negotiated by the CC functionality.

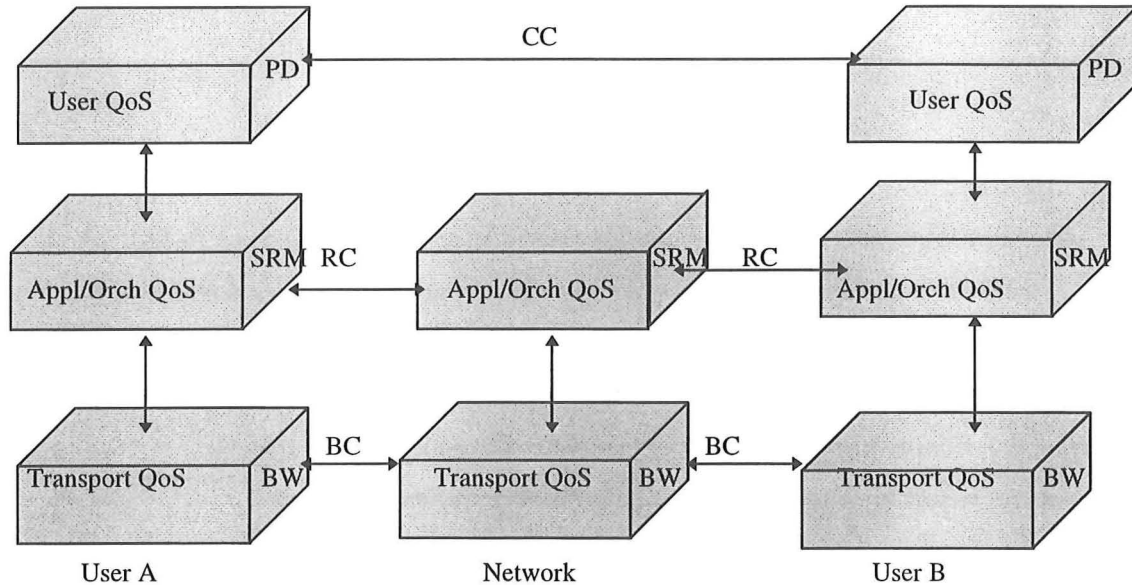


Figure 3.1 : Resource Allocation

3.2.1 Simultaneous allocation of Resources

In this situation, the resources are allocated concurrently with end-to-end call setup and hence, the call setup time is low for this type of allocation. In POTS or N-ISDN, simultaneous allocation was the primary method used for resource allocation. This was mainly due to the monolithic signalling protocol that existed in those services. In providing POTS services over B-ISDN, simultaneous allocation will have to be used to simulate the POTS environment [RACE MAGIC deliverable 14]. From the customer point of view simultaneous allocation would provide a high set up Quality of Service, because of the low call setup time.

3.2.2 Post Allocation of Resources

In post allocation, the end-to-end call establishment is carried out before the resources are allocated. One obvious disadvantage of this method is, that the participants have to wait until the resources are allocated after establishing the call. A user may want to raise or lower the QoS of a bearer connection after information transfer begins. Also a user may want to lower the QoS of one bearer connection to raise the QoS of another. For example, the user may lower the video resolution from CIF to QCIF in order to raise the audio QoS from

G.722 to Hi-fi. In the above scenarios, post allocation is the only possible method of resource allocation. This method of resource allocation is also called sequential setup. The terms, end-to-end call establishment delay, special resource establishment delay and bearer connection establishment delay may be defined for this type of allocation. Call setup time may be high in this case and this results in a lower setup QoS. In case of multipoint connections, post allocation is more feasible as the capabilities of end terminals may be negotiated before valuable special resources are allocated.

3.3 Code Converters

Code converters are used to convert isochronous streams such as audio and video from one format to another. Figure 3.2 is a symbolic representation of a code converter. In audio the code conversion would be from one audio format to another. The possible code conversions in video are illustrated in table 3.1. The code conversion function may be performed at the receiver if the necessary hardware is available. Otherwise an intermediate code converter may have to be used. If an intermediate code converter is used, three clocks (sender clock, receiver clock and the code-converter clock) are involved, and the jitter in the end-to-end delay is high due to there being more slip/pause stages. The reader is referred to chapter 4, for a description of the slip/pause mechanism.

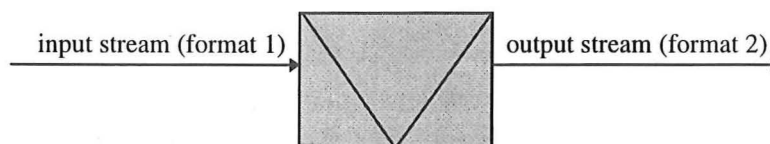


Figure 3.2 : Code Converter

Type of conversion	Example
conversion of video standard	MPEG-1 to H.261
frame rate conversion	30 frames per second (fps) to 10 fps
frame size conversion	CIF to QCIF
colour to black and white conversion	colour NTSC to black & white NTSC
conversion of number of colours	512 colours to 256 colours
conversion of shades of grey	256 shades of grey to 128 shades of grey
aspect ratio conversion	4:3 to 16 :9

Table 3.1 : Video Code Conversion

3.4 Multicasting Switches

POTS and N-ISDN switches were only capable of unicast communications. In order to set up a multicast connection the users had to be connected in a mesh network. However, in B-ISDN, ATM switches which are capable of multicast communications are being introduced. Figure 3.3 gives the symbolic representation of a multicast switch. In these switches, in addition to a switch fabric which does the ATM cell switching function, a copy fabric is also incorporated in order to provide multicast capability. The copying function copies cells from one incoming channel to any subset of its outgoing channels [Lee, 1988]. In specialized applications such as network games, multicasting switches may be used to send a stream from a particular source to multiple destinations and these streams may have to be played in synchrony in all the destinations [Escobar *et. al.*, 1992]. The multicasting switch adds almost negligible delay or jitter to media streams.

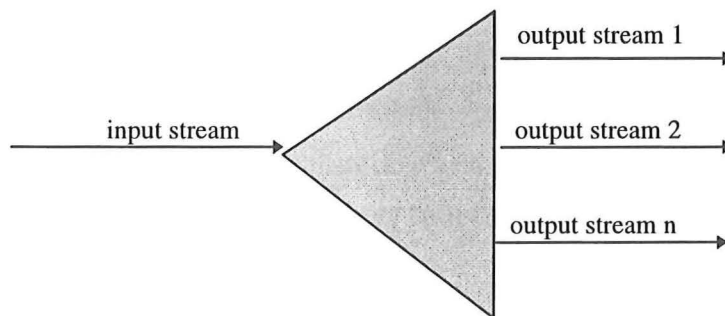


Figure 3.3 : Multicasting Switch

3.5 Encoders and Decoders

In this section, the audio / video encoders and decoders pertaining to different standards are described. In section 3.6, audio and video combining techniques are presented.

3.5.1 Audio Encoders and Decoders

Table 3.1 summarizes the audio encoding standards. ADPCM coding generates a lower bit rate in comparison to PCM due to the efficiency of coding. Thus for a telephone quality signal of 0.3 - 3.4 KHz, G. 711 (PCM), generates a higher bit rate of 64 Kbits/s in comparison to G.721 (ADPCM) which only generates a bit rate of 32 Kbits/s. Also G.722 (ADPCM) only generates a bit rate of 64 Kbits/s for a wideband signal of 0.5 - 7 KHz.

G. 728 uses Low Delay - Code Exited Linear Prediction (LD-CELP) for encoding. This encoding method generates a highly compressed high quality signal although at a high computational complexity. Hi-fidelity codecs are needed for coding audio applications such as music, which require the full audio spectrum.

Standard	Bit Rate (Mbits/s)	Bandwidth (KHz)	Coding Method
G.711	0.064	0.3 - 3.4	PCM
G.721	0.032	0.3 - 3.4	ADPCM
G.722	0.064	0.05 - 7	ADPCM
G.728	0.016	0.3 - 3.4	LD-CELP
Hi-fidelity	0.192 - 1.411	0.02 - 20	not standardized

Table 3.1 : Audio Standards

A voice stream consists of talkspurts and silence periods. A technique known as Digital Speech Interpolation (DSI) is used to suppress the silence periods so that better compression is achieved. However this produces bursty voice sources [Sriram *et. al.*,1991], and as a result Variable Bit Rate (VBR) streams are produced. The VBR traffic could introduce congestion during bursty periods. A prioritization scheme is used to prevent loss of important information during congestion. Here Most Significant Bits (MSBs) of the voice stream are tagged as High Priority (HP) cells while the Least Significant Bits (LSBs) are tagged as Low Priority (LP) cells. In times of congestion, LP cells are dropped in preference to HP cells, and overall loss in QoS is minimized.

3.5.2 Video Encoders and Decoders

3.5.2.1 H.261

H.261 was the earliest digital video standard to be introduced for motion video. The H.261 picture is divided into luminance blocks of size 8x8 pixels. Four luminance blocks are included in a macroblock. Figure 3.4 depicts a block diagram of the H.261 coding process. There are two types of frames in H.261. These are called intra-frames (I frames) and inter-frames or predictive frames (P frames). In the figure, intra-frames are transmitted if the selector gets the choice 0. Intra-frames are DCT coded and quantized before being entropy coded (VLC coded). The first frame and a frame at a scene cut is intra-frame coded. In addition at least, every 132nd frame is intra-frame coded.

compared to H.261. MPEG-4 is designed to address the need for a variety of applications where either the channel has a low bit rate or the storage medium has a reduced capacity [Chiariglione, 1995]. MPEG standards utilize procedures similar to H.261 to code the frames. However in the MPEG standard, bidirectional temporally predicted frames (B frames) are also used.

Figure 3.5 depicts a sequence of frames pertaining to the MPEG standard. The number of bits of a B frame, is generally lower than both I frames and P frames. Generation of B frames, requires more processing overhead. The bidirectionally predicted frames incur a frame reordering delay at both the encoder and the decoder. This delay can be objectionable for interactive communications.

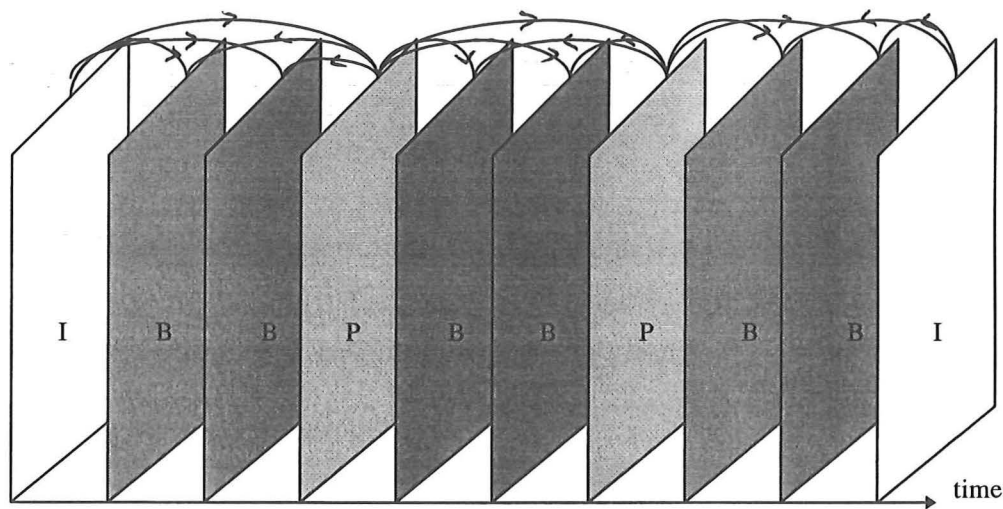


Figure 3.5 : MPEG sequence of I,P,and B frames

MPEG sources produce highly bursty VBR traffic, as intra-frames can produce a very high bit rate. Thus a prioritization scheme is essential in transmitting MPEG information across the network, to limit the QoS degradation under congested periods. All header information, motion vectors, main DCT coefficient (DC value) and a few other DCT coefficients (AC values) may be transmitted as high priority information while the remaining DCT coefficients may be transmitted as low priority information [Pancha *et. al.*, 1994].

The Joint Photographic Expert Group (JPEG) Standard is defined for the transmission of high quality still pictures. Transmission of video consisting only of high quality intra-frames is called motion JPEG [Sae Zon-Yin *et. al.*, 1993; Bauer *et. al.*, 1995]. Motion JPEG sources are not bursty but produce a very high bit rate.

3.6 Media Combiners

Media combiners are used to combine two or more streams together. Media combiners are used only for isochronous services such as audio and video. The type of combining used in audio is very much different to that of video. A Concast Switch is another name given for a media combiner. The figure 3.6 gives a symbolic representation of a media combiner.

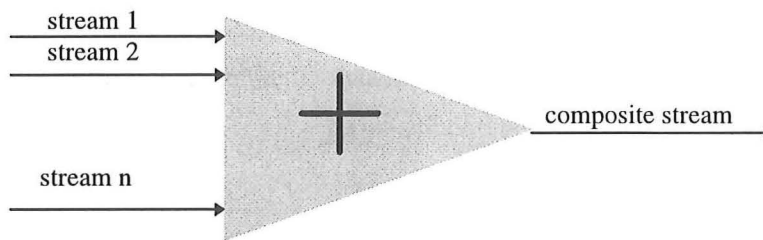


Figure 3.6 : Media Combiner

3.6.1 Combining Audio

Audio mixing is the primary operation performed in combining audio. The different audio streams to be combined are multiplied by individual gain factors and then added together. The mixing operation is performed either in the digital domain or the analog domain. Figure 3.7 represents the mixing operation.

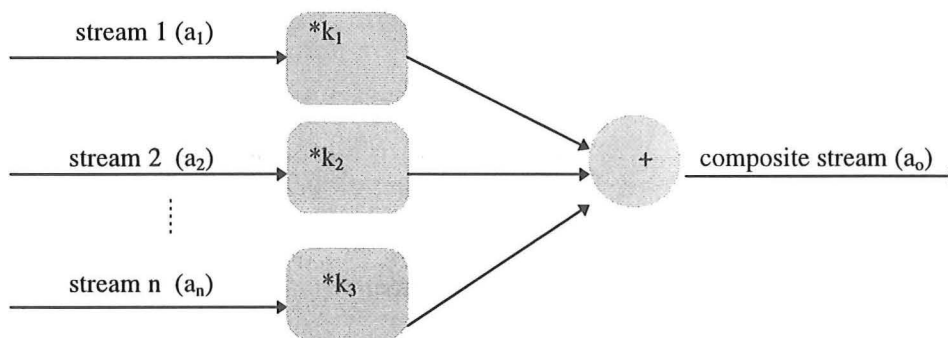


Figure 3.7 : Audio Mixer

If the input and output streams are represented as shown in the diagram, the output stream is given by $a_o = k_1 * a_1 + k_2 * a_2 + \dots + k_n * a_n$. The above process is non-reversible, as the original streams cannot be reconstituted from the composite stream. Time Division Multiplexing (TDM) and Frequency Division Multiplexing (FDM) methods used in POTS, can also be considered as audio combining methods. If the combining process is carried out in the digital domain, the processing delays would be lower compared to that of the analog domain.

3.6.2 Combining Video

Especially in video conferencing applications there is a need to view several video sources as separate windows on the screen. This situation is also called *picture-in-picture* [Singh *et. al.*, 1993]. Figure 3.8 depicts one possible scenario where four QCIF windows may be combined to produce one CIF image.

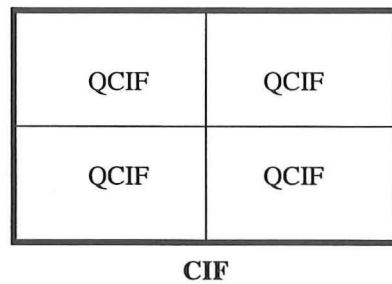


Figure 3.8 : Combining four QCIF windows to produce one CIF image

Three different approaches are suggested for combining video, namely

- pel domain approach
- DCT domain approach
- VLC domain approach

3.6.2.1 Pel Domain Approach

In the pel domain approach, the video data streams are fully decoded, combined and encoded again. This approach has the highest flexibility as different coding standards and algorithms may be accommodated. However the processing delays will be high and the overall picture quality may be low due to encoder and decoder stages. Figure 3.9 depicts a generalized pel domain video combiner.

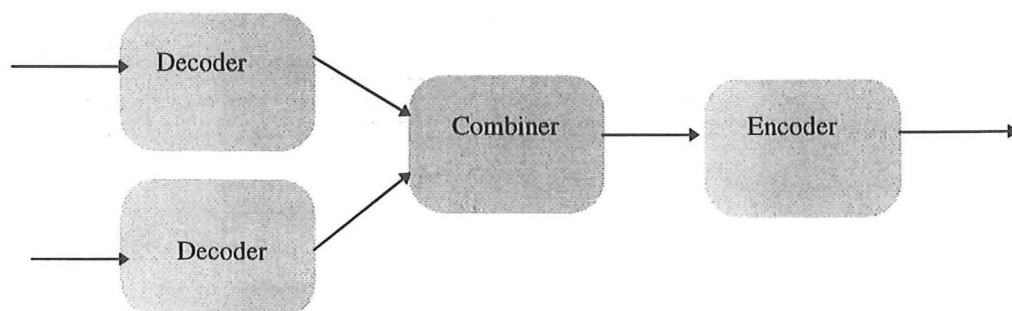


Figure 3.9 : Generalised pel-domain video combiner

3.6.2.2 DCT Domain Approach

DCT domain approach is a method suggested by S-F Chang *et. al.*[1993]. Conversion of all video streams down to the pel domain to combine them incurs additional delays and expenses. In this method, only the variable length codes are decoded and the combining is done on DCT coefficients in the spatial blocks. This method can be applied even in cases where the block boundaries of the input and output images change. Consider the situation depicted in figure 3.10 and figure 3.11.

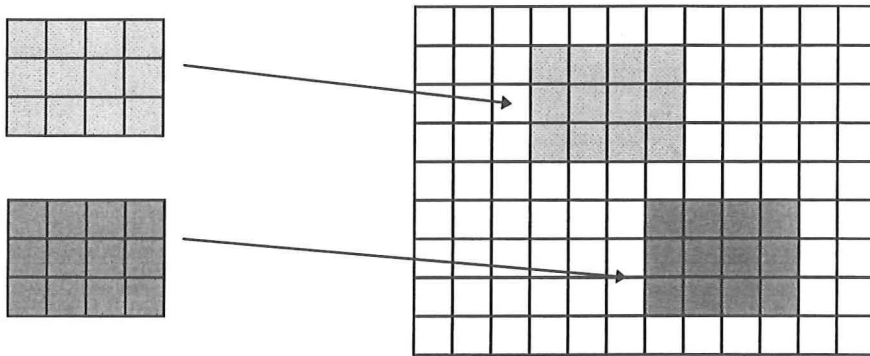


Figure 3.10 : Perfect Matching of input and output DCT blocks

In figure 3.10, there is a perfect match of input and output DCT blocks. However, in figure 3.11, there is an overlap of input and output blocks as block P has to be constructed from four adjacent input blocks. Therefore, video combining cannot be performed in the fully encoded domain and full decoding into the pel domain is necessary to calculate the DCT coefficients of the output block. In the approach suggested by S-F Chang *et. al.*[1993], transition matrices are defined to calculate the DCT coefficients of the output block from the DCT coefficients of the input blocks.

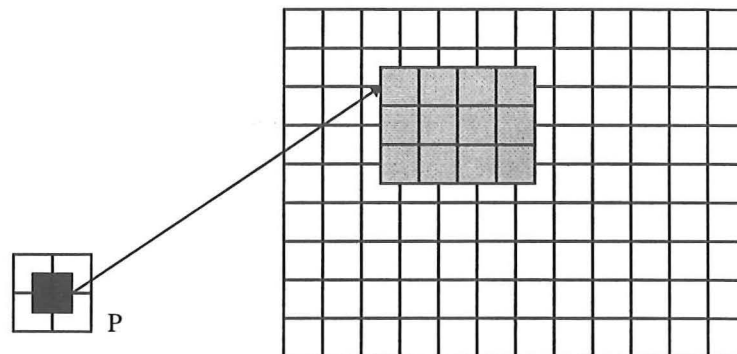


Figure 3.11 : Overlapping of input and output DCT blocks

3.6.2.3 VLC Domain Approach

S-M Lei *et. al.*, suggests a combiner that processes H.261 data headers of four QCIF streams and concatenates the remaining data streams without modification to produce one CIF image. They refer to this method as the VLC domain approach as data coded in Variable Length Codes (VLCs) is processed. The primary task of the combiner is to multiplex variable length coded Group Of Blocks (GOB) into a single bit stream. The processor detects picture headers and GOB headers and then relabels these headers. The picture header of first QCIF stream is retained and its format is converted from QCIF to CIF and all the other QCIF headers are discarded. The Group Numbers (GNs) of each QCIF is changed as shown in figure 3.12. The VLC domain approach is only valid for situations where the block boundaries do not change.

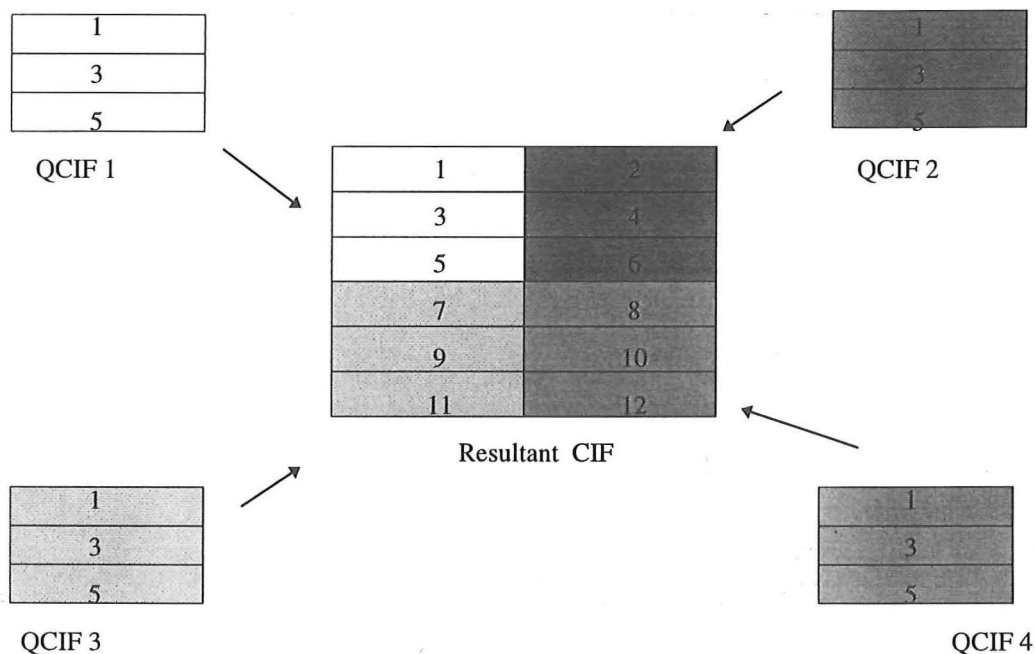


Figure 3.12 : Video Combining in VLC domain

3.6.3 Synchronization of Combined Streams

In multimedia conferencing applications audio and video originating from two or more sources may be combined before presentation. The audio and video of each source has to be lip-synchronized at playout. The combining process may be performed at an intermediate location or at the receiver. If the combining process is performed at an intermediate location, the jitter in the end-to-end delay of the media stream is high due to the necessity of three slip/pause stages. If more media sources are combined there will be more synchronization constraints to be satisfied, in combining the streams. Example 7.3 illustrates how audio and

video streams could be synchronized using Static Delay Compensation mechanism proposed in this thesis.

3.7 Layered Encoders and Decoders

Layered encoders and decoders are used in isochronous services such as audio and video. Layered encoding is the process of separating the original signal into a few constituent layers. The separation process produces a base layer which contains the primary information and a few other hierarchical layers. The addition of each layer would enhance the quality of the signal. Layered encoders and decoders are symbolically represented by figure 3.13. Layered coding has the following advantages over non layered coding [Coppisetti *et. al.*, 1993, Singh *et. al.*, 1993, Shacham *et. al.*, 1992, Lim *et. al.*, 1992].

- Analysis and synthesis for a particular layer is simpler than that for the overall signal.
- Effect of errors in the overall image after reconstruction is small because an error occurring in decoding a signal in a particular layer is independent of the other layers.
- Layers may be quantized using different step sizes thereby controlling the number of bits allocated to each layer.
- Parallel processing techniques may be applied.
- This facilitates the creation of encoded bit stream that may be decoded in part if the channel bandwidth drops, decoder resources are limited or a smaller image than the source is desired.
- If a particular media is to be transmitted to multiple destinations, the different receivers could decode one or more layers depending on their decoder capabilities and requirements.

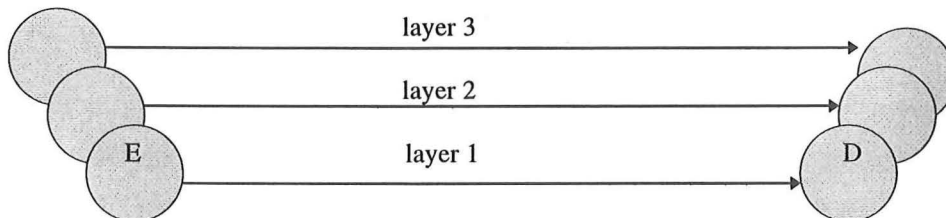


Figure 3.13: Layered Encoder and a Decoder

Spatial and temporal frequency band decomposition are the techniques being used for layered coding of video. Layered coding schemes are categorized according to the type of

filter being used for the separation of layers. The resulting coding process is called subband coding or wavelet coding depending on whether quadrature mirror filters or wavelet filters, respectively, are being used [Zafar *et al.*, 1993]. The remainder of this section discusses the different layered coding techniques in detail.

3.7.1 Spatial Band Decomposition

Spatial band decomposition is the process of separating a video signal according to horizontal and vertical frequencies. Since the sensitivity of the Human Visual System (HVS) is higher at the low spatial frequencies, the baseband is chosen to be the one with low horizontal and vertical frequencies. Figure 3.14 shows the different layers of a two layer subband coding system. Short Kernel Filters (SKF) which are a brand of Quadrature Mirror Filters are used for subband coding in the method suggested by Coppisetti *et al.*[1993].

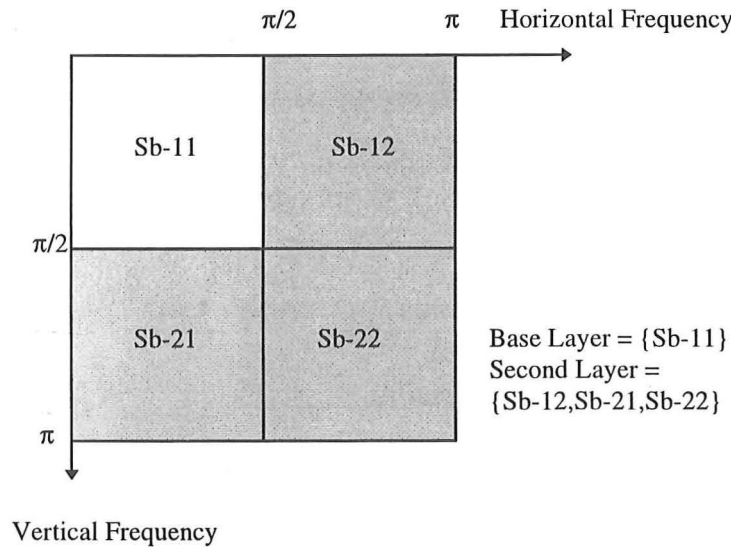


Figure 3.14 :Two Layered Subband Coding

Usage of SKF on subband decomposition is depicted in figure 3.15. If the high pass filtering and the low pass filtering is continued to another stage the subbands as shown in figure 3.16 would result. The spatially decomposed subbands may be compressed using different quantizer step sizes. High frequency subbands are subjected to high compression while low frequency are subjected to low compression. The quantizer step size can also be varied on a block by block basis.

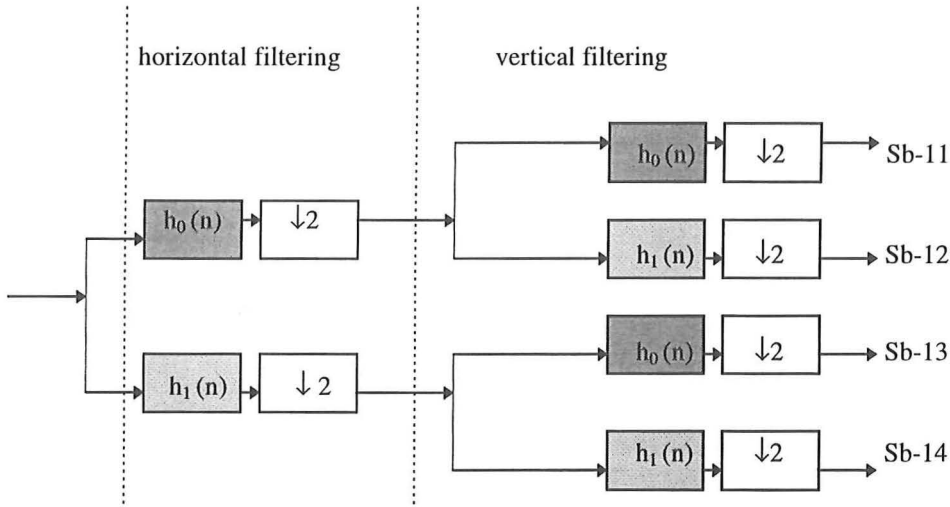


Figure 3.15 :Two Layered Subband Coding by low pass filters ($h_0(n)$) and high pass filters ($h_1(n)$), where $\downarrow 2$ indicates downsampling

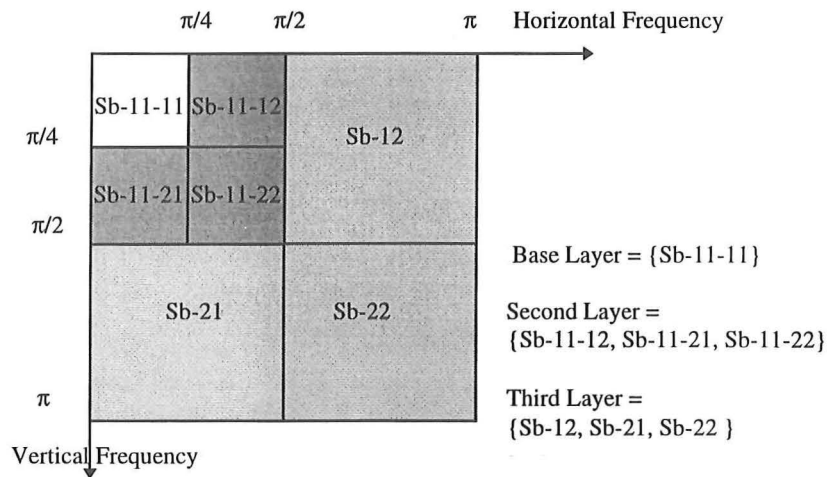


Figure 3.16 :Three Layered Subband Coding

3.7.2 Temporal Band Decomposition

Temporal band decomposition is the process of separating the 30 Hz temporal frequency band into a few temporal bands. M-band filter sets can be used to decompose the original signal into separate temporal bands [Singh *et. al.*, 1993]. Temporal coding may also be described as motion sensitive coding where areas of an image with different motion values are coded differently. Sometimes a combination of spatial and temporal coding (spatio-temporal decomposition) is used to produce different layers of a layered coding system.

3.7.3 Spatio-Temporal Decomposition

In the spatio-temporal decomposition process proposed by Hollier *et. al.* [1990], a difference frame is produced by subtracting the current frame from the previous frame. This difference frame is divided into a baseband and a higher band, which relate to base layer and second layer respectively, in figure 3.14. The macroblocks of the baseband are divided into four motion groups depending on the motion activity in each of them. Every macroblock of the baseband has a related macroblock in the higher band. Each macroblock in the higher band, is assigned to one of four motion groups according to the motion group its corresponding baseband macroblock belongs. Macroblocks of all four motion groups in the higher band are further divided into macroblocks with high or low detail. This produces twelve groups of macroblocks with eight in the higher band and four in the baseband. Any combination of these bands may be used to produce a number of layers. Spatio-temporal decomposition can be used to separate video so that the resulting subbands could be coded according to the Human Visual Characteristics (HVS).

3.7.4 Layered Coding Based on Human Visual Characteristics

There has been a lot of research done in layering the video signals according to the Human Visual System [Kim *et. al.*, 1993]. The Human Visual System is more sensitive to low horizontal and vertical bandwidths and hence the base layers of two layer and three layer systems are coded with low compression. In a three layer system, layer 2 is coded with medium compression and layer 3 will be coded with high compression. In the Human Visual System, the sensitivity to the motion is also reduced with increasing horizontal or vertical frequencies. Singh *et.al.* [1993], suggested a spatio-temporal decomposition method based on HVS for layered coding of video. In this method, the video frames are divided into twenty bands as shown figure 3.17.

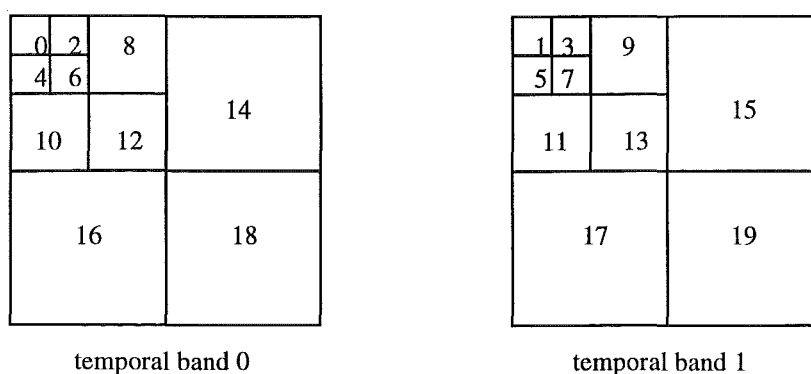


Figure 3.17 : Spatio-temporal coding based on HVS

Temporal band 0 relates to low temporal frequency while temporal band 1 relates to high temporal frequency. Subjective evaluations of reconstructed images had been made by assigning varying bit rates for different subbands. It is evident from the experiments reported in the paper that, an acceptable quality image could be produced by using the bands 0-11 and the bands 15, 17 and 19 may be left out without any perceivable quality degradation.

3.7.5 Synchronization of Layered Coding Systems

Layered coding is a block based coding process, and encoder and decoder processing delays of the different layers are equal. However, for optimized bandwidth allocation, the layers may be transmitted across different paths and hence their transfer delays could be different. As the base layer contains the essential information it is transmitted over a high quality path with low cell loss characteristics[Shacham, 1992]. Higher layers are able to tolerate high cell losses and therefore, they may be transmitted over a path with moderate cell loss characteristics.

Consider a multicast layered video configuration as shown in figure 3.18. At the sender, the video stream is decomposed into two layers. Assume that, the terminal equipment of receiver B is only capable of decoding layer 1 while terminal equipment of receiver C is capable of decoding both layer 1 and layer 2. Layer 1 may be sent along a different path to layer 2, for optimized resource allocation. At receiver C, layer 1 and 2 need to be synchronized before presentation. In chapter 7, a mechanism is proposed by the author to synchronize the layers of a layered coding system.

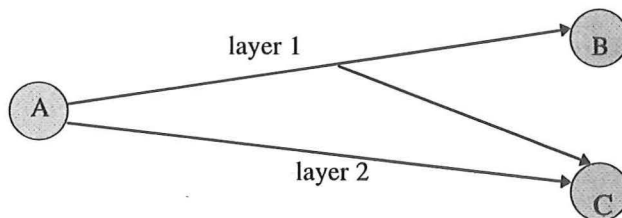


Figure 3.18 : Multicast layered video configuration

If layered encoding is used to encode a stream, at the receiver only the layers pertaining to a particular frame can be decoded together, i.e. if one layer belongs to a particular frame, and the other layer belongs to an adjacent frame, they cannot be decoded together.

Synchronization of layered coding systems is further discussed in section 7.4.

3.8 Media bridges

Media bridges can be considered as centralized pieces of equipment which are a combination of other special resources such as code converters, multicasting switches and media combiners. These are used for multimedia conferences. A Multipoint Control Unit (MCU) is another name given for a media bridge [Willebeek-Lemair *et.al.*, 1994]. Figure 3.19 represents a media bridge. Each of the sinks would receive a combination of inputs from other sources. Video received by sink 1 may be the combined output of sources 2 and 3. Thus services of a media combiner may be utilized. Also if the video formats of source 2 and sink 1 differs, a code converter has to be utilized. If sink 2 and sink 3 requires source 1 audio then a multicasting switch has to be used as well.

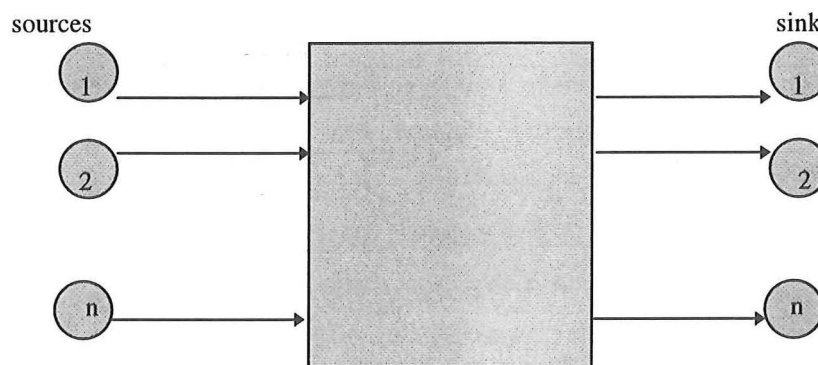


Figure 3.19 : Media bridge

The MCUs used in N-ISDN provided for full audio combining but only allowed selection of a particular video stream from multiple video streams. However in B-ISDN, a media bridge will have to incorporate a lot of other special resource functions and its value would increase with the versatility it provides. Intelligent media bridges can be equipped with protocol functionality to allocate different special resources and establish synchronized bearer connections, when setting up multimedia calls.

3.9 Summary

In chapters 2 and 3, the QoS parameters that need to be guaranteed in connecting multimedia calls in Broadband ISDN were analysed. This included the analysis of audio, video and synchronization QoS parameters associated with different special resources. In the next few chapters a mechanism for guaranteeing synchronization QoS of multimedia calls is proposed and analysed. In chapter 4, an intra-media synchronization mechanism is proposed to ensure playout continuity of any isochronous media stream. In chapters 5,6 and 7 an

inter-media synchronization mechanism called Static Delay Compensation (SDC) is proposed to synchronize any number of media streams.

CHAPTER 4

SYNCHRONIZATION

Synchronization is a very important aspect of QoS management. It includes both intra-media synchronization and inter-media synchronization. In this chapter, a mechanism is developed for intra-media synchronization, taking into account the jitter in the end-to-end delay of a media stream and also lack of synchrony in sender and receiver clocks. The different delays that contribute to the end-to-end delay of a media stream are analysed and robust a priori estimates for the minimum and maximum bounds of delay are derived. This robust delay behaviour can be utilized in facilitating inter-media synchronization.

4.1 Sender and Receiver Clocks

Media frames are produced at the sender and presented by the receiver to the user. Figure 4.1 depicts the different processes involved at both the sender and the receiver.

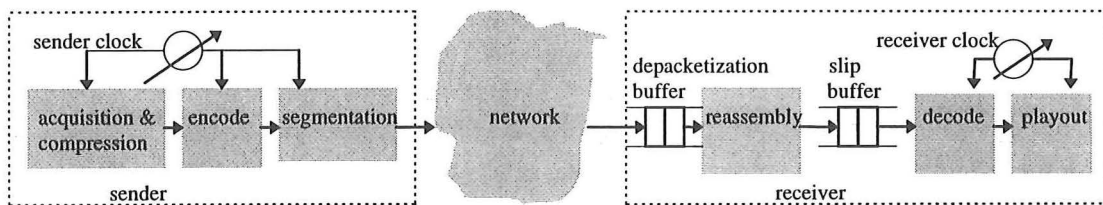


Figure 4.1 : Processes at the sender and the receiver

A service clock at the sender determines the rate at which media frames are acquired, compressed, encoded and segmented for transmission. The receiver's service clock determines the rate at which the media frames are re-assembled, decoded and played out. The depacketization buffer stores the ATM cells to be re-assembled while the slip buffer stores media frames to be decoded. In conventional analog television receivers as well as in circuit switched connections, the receiver clock is derived by phase locking onto the service clock from the sender. The above phenomenon is called *connection-driven timing* [Anderson *et. al.*, 1991]. In packet switched networks, to implement *connection-driven timing*, time

stamps are sent along with picture information so as to emulate the circuit switched system. However, due to the jitter in the end-to-end delay in packet switched networks, the above method does not work perfectly [Lau *et. al.*, 1992].

In packet switched networks, the display timing at playout cannot be derived accurately from the sender's service due to the jitter. Thus, display timing should run independently from the sender's clock. The decoder should be fully synchronized with display timing to prevent the *horizontal tear effect* [Katseff *et. al.*, 1991]. Thus, the decoding and display processes at the receiver, are controlled by the receiver's clock, independently from the sender's clock, as shown in figure 4.1. This phenomenon is called *device-driven timing*. However for continuous operation, the receiver's clock should have no drift with respect to the sender's service clock. If the sender's service clock runs faster, the receiver will not be able to cope with the arrival of frames (buffer overflow) and if the receiver clock runs faster then the receiver will be starved of frames (buffer underflow).

4.1.1 Clock Recovery Methods

A clock recovery method may be used to synchronize the sender and receiver clocks to prevent buffer underflow or buffer overflow. Synchronous Residual Time Stamp (SRTS) and adaptive clock recovery method are two methods suggested for clock recovery.

4.1.1.1 Synchronous Residual Time Stamp (SRTS)

In the SRTS [CCITT, 1991] method, the sender and receiver clocks are both calibrated against a common network reference clock. The sender sends information to the receiver that convey the frequency difference between its service clock and the reference clock, from which the receiver can adjust its service clock. This method converges faster than the adaptive clock recovery method but can only be implemented in a network with a common reference clock. SRTS method is expensive to implement due to the requirement of a network reference clock.

4.1.1.2 Adaptive Clock Recovery Method

In the adaptive clock recovery method suggested by De Prycker *et. al.*[1987], Singh *et. al.* [1988] and Ahmed [1989], the rate of filling of the depacketization buffer at the receiver is used to adjust the receiver's service clock. Over a large interval, fluctuations in the buffer

filling rate due to cell delay-jitter is zero but any frequency offset between the sender and receiver clocks causes the buffer filling rate to steadily increase or decrease. By monitoring the buffer filling rate over a sufficiently large interval, an accurate estimate can be made of the sender-to-receiver frequency offset, and the receiver clock can then be adjusted to match the sender's service clock. In adaptive clock recovery method, as the buffer level is monitored over large intervals, the convergence time for adjustment of the receiver's service clock can be large.

4.1.2 Slip and Pause

Clock synchronization methods can be used to minimize the drift of the receiver's clock with respect to the sender's clock. Clock recovery methods require protocol processing at the end terminals and transfer of control messages across the network. However, perfect synchronization between the sender and receiver clocks may not be achieved due to errors in the frequency measurement and also sudden temperature variations of crystal oscillators driving the clocks.

The sender to receiver clock asynchrony is bounded because the crystal oscillators driving these clocks are manufactured according to bounded frequency deviation specifications [De Prycker *et. al.*, 1987]. A slip/ pause mechanism may be used to ensure media continuity if sender to receiver clock asynchrony is bounded. The slip/ pause mechanism, eliminates the need for an expensive clock recovery method. The receiver can either slip (leave out) the display of a media frame to decrease the *slip delay* or pause (repeat) the display of a media frame to increase the *slip delay*. The slip delay is the time that each data unit is stored within the slip buffer.

The slip delay is monitored to determine when media frames should be slipped or paused, to alleviate any overflow or underflow of the slip buffer. The upper threshold of the slip delay must be set so that when a media sample is slipped it is impossible for the data unit or frame, immediately following to have a slip delay below the lower threshold, since that can cause slips and pauses to alternate rapidly. This method was first mooted by Cochenne *et. al.*, [1985]. How the slip/pause mechanism can be used to ensure intra-media synchronization, is described in section 4.3. The effect of transfer delay-jitter on the slip/pause mechanism is elaborated in section 4.5 and examples on slip/pause frequency are given in section 4.6.

4.2 End-to-end Delay of a Media stream

Any media stream that is transmitted across a digital network has to undergo different delays during its production, transmission and consumption. These delays could be categorized as

- processing delays (δ^{enc} , δ^{dec})
- transfer delay (d)
- slip delay (Δ)

Figure 4.2 depicts these delays and the shaded regions represent the variable parts of the delays. The processing delays could be further divided into encoder processing delays (δ^{enc}) and decoder processing delays (δ^{dec}). The transfer delay consists of the segmentation and reassembly delay, propagation delay across the network, and queuing delay at ATM switches. The maximum transfer delay is denoted by d^{max} , and the maximum transfer delay-jitter is denoted by j^{max} . As the minimum transfer delay can be obtained by subtracting the maximum transfer delay by the maximum transfer delay-jitter, the minimum transfer delay is given by ($d^{max} - j^{max}$).

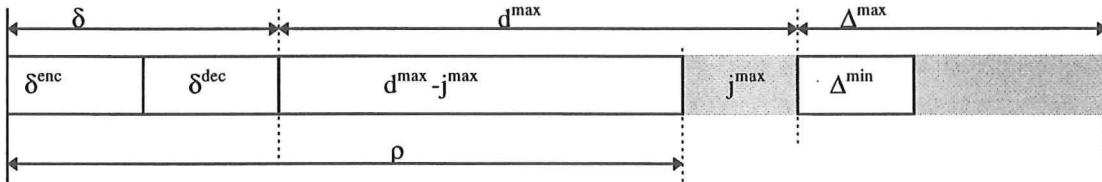


Figure 4.2 : End-to-end delay of a media stream

The minimum and maximum values of the slip delay is denoted by Δ^{min} and Δ^{max} respectively. The value ρ denotes the minimum elapsed time of any media frame from the time of its generation to the time it arrives at the slip buffer. Therefore the value ρ is equal to the sum ($\delta^{enc} + \delta^{dec} + d^{max} - j^{max}$). The next section derives values for the minimum and maximum slip delay to ensure intra-media synchronization.

4.3 Intra-media Synchronization

Intra-media synchronization is a mechanism that is introduced to counteract the effects produced by both sender and receiver clock asynchrony as well as the delay-jitter introduced

by the network. The minimum elapsed time of any media frame from the time of generation to the time of arrival at the slip buffer is given by ρ , as shown in section 4.2. Figure 4.3 shows the arrival times of the media frames at the slip buffer and the times at which media frames are displayed.

If media frame 1 is acquired at time zero at the sender it will arrive at the slip buffer at a minimum and maximum time of ρ and $\rho + j^{\max}$ respectively. Media frame 2 will arrive at the slip buffer at a minimum and maximum real time of $\rho + \tau$ and $\rho + \tau + j^{\max}$ respectively, where the value τ represents the frame duration. Ideally the receiver clock also occurs at intervals of τ . At clock occurrence, Clk 1 the decoder processes frame 1 and displays it until the occurrence of Clk 2. If there is no clock asynchrony between sender and receiver clocks, frames can be displayed continuously as shown in figure 4.3.

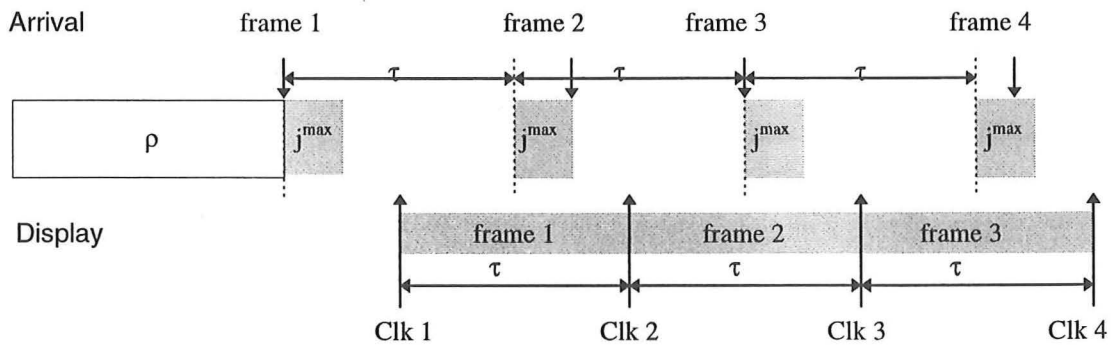


Figure 4.3 : Arrival times of media frames to the slip buffer and display of frames

However, if the receiver clock lags behind the sender clock, after a while the situation depicted in figure 4.4 could occur. Frame n is displayed between the interval Clk n and Clk $n+1$ as frame n arrives before the clock occurrence Clk n . However as frame $n+1$ arrives after the clock occurrence Clk $n+1$, media frame n is re-displayed (paused) between clock occurrences Clk $n+1$ and Clk $n+2$. Frame $n+1$ is buffered and only displayed during the interval between Clk $n+2$ and Clk $n+3$.

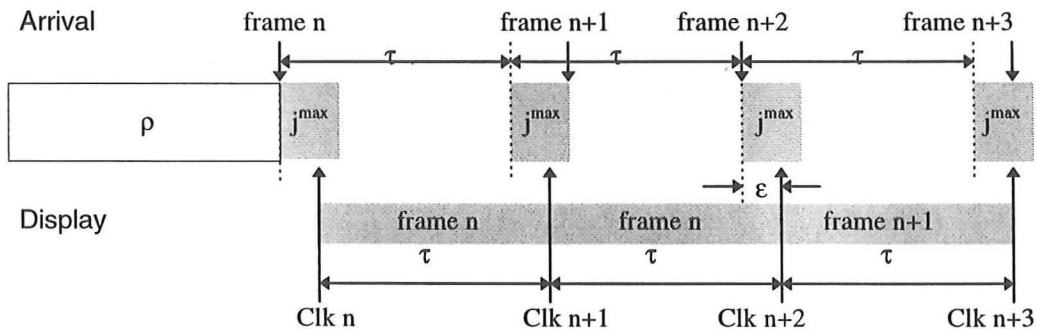


Figure 4.4 : Pausing a media frame

The slip delay of frame $n+2$ is equal to $\tau + \epsilon$. As ϵ is always less than j^{\max} , slip delay of frame $n+2$ could never be greater than $\tau + j^{\max}$. If the maximum slip delay of a media frame (Δ^{\max}) is made equal to $\tau + j^{\max}$ a media frame slip will not occur immediately after a media frame pause. The minimum slip delay is zero as the slip delay cannot be negative. Thus to ensure intra-media synchronization, minimum and maximum slip delays are made equal to 0 and $\tau + j^{\max}$ respectively.

Figure 4.5 depicts the end-to-end delay of a media stream after the minimum and maximum slip delays have been substituted in figure 4.2. The term J^{\max} represents the maximum jitter of the total end-to-end delay of the media stream and it is equivalent to $\tau + 2j^{\max}$. The term σ is the maximum end-to-end delay of the media stream and is equivalent to $\rho + \tau + 2j^{\max}$.

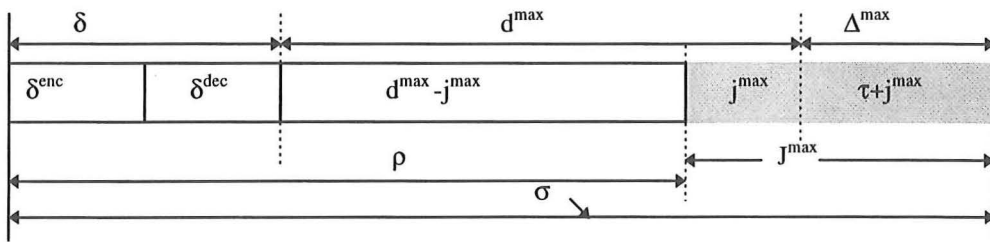


Figure 4.5 : End-to-end delay of a media stream (minimum and maximum slip delays being substituted)

4.4 Examples of Total Jitter in End-to-end Delay

Consider a H.261 video stream as well as a G.722 audio stream transmitted across the Broadband Integrated Services Digital Network

H.261 video frame size (τ) = 33 ms,

If the stream has to go through two ATM switching stages and if each stage produces 250 μ s of maximum jitter, $j^{\max} = (250 * 2) \mu\text{s} = 0.5 \text{ ms}$

Therefore the total end-to-end delay-jitter (J^{\max}) = $\tau + 2 * j^{\max} = 34 \text{ ms}$.

Similarly G.722 audio frame size (τ) = 5.75 ms,

If the stream has to go through two ATM switching stages, $j^{\max} = 0.5 \text{ ms}$

Therefore the total end-to-end delay-jitter (J^{\max}) = 6.75 ms.

If satellite communications were used instead of fibre optic communications, each satellite hop would introduce about 40 ms additional delay-jitter. If two ATM switching stages and a single satellite link is used for the communication

total jitter of the end-to-end delay of video (J^{\max}) = $\tau + 2*j^{\max} = 33 + 2(40+0.5) = 114$ ms

total jitter of the end-to-end delay of audio (J^{\max}) = $\tau + 2*j^{\max} = 5.75 + 2(40+0.5) = 86.75$ ms

These values are objectionable for audio and video synchronization as shown in section 5.3.

4.5 Effect of Jitter on Slip /pause Mechanism

If the receiver clock is slower than the sender clock, after a while a media frame may have to be slipped. There is no exact position in time where the slip could occur, due to the jitter in the transfer delay (j^{\max}). The next subsection illustrate the extremes of the earliest slip and the latest slip as well as an intermediate slip example. If the receiver clock is faster than the sender clock, a media frame may have to be paused. The effect of jitter on a media frame pause is illustrated in subsection 4.5.2.

4.5.1 Slip

As shown in figure 4.6, at the occurrence of clk n+1, the slip delay of media frame n+1 is equal to $\tau + j^{\max}$. As the maximum allowable slip delay of a media frame is $\tau + j^{\max}$, media frame n+1 is slipped, and media frame n+2 is displayed between the interval clk n+1 and clk n+2. The slip occurs at the earliest position as the frame arrived at the earliest time and the clock occurred after a time of $\tau + j^{\max}$ as shown in the diagram.

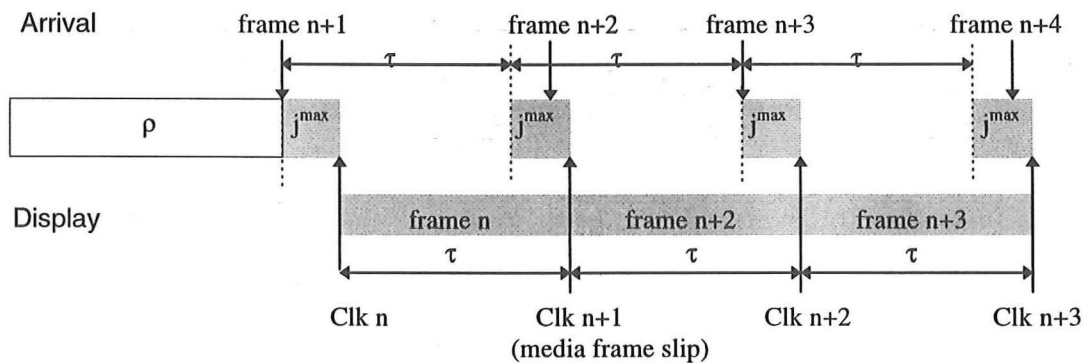


Figure 4.6 : Slipping a media frame with elapsed time p

When the network is heavily loaded, media frames can get delayed continuously, and a media frame slip will not occur until the clock has drifted further. In figure 4.7, the clock has drifted further by ζ , and the media frame $m+1$, that arrives at the slip buffer at an elapsed time, less than $\rho + \zeta$ is slipped. Note that, even if a single frame arrives at an elapsed time less than $\rho + \zeta$, a media frame will be slipped.

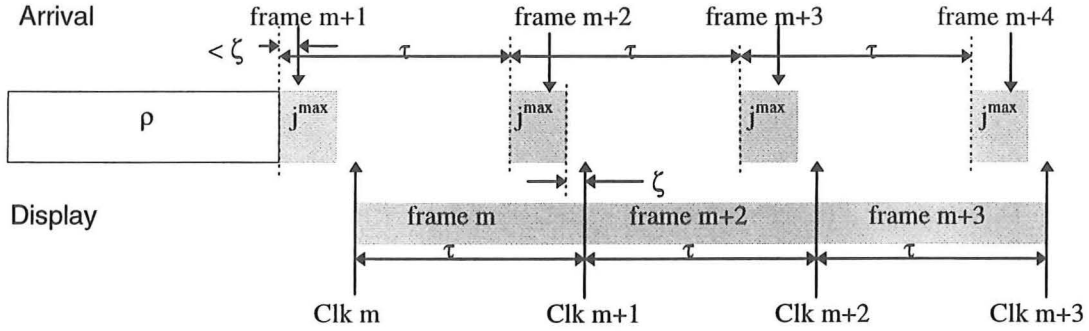


Figure 4.7 : Slipping a media frame with elapsed time $\rho + \zeta$

If the frames arrive at the latest possible elapsed time of $\rho + j^{\max}$, continuously, a media frame slip will not occur until the situation shown in figure 4.8. Thus, one cannot predict exactly the time at which the slip or pause can occur. The time interval where this uncertainty prevails is equivalent to the time taken for the receiver clock to drift by j^{\max} . This time interval is referred to, in this thesis, as the *uncertainty region*. The phenomenon of uncertainty in the slip/pause mechanism affects the synchronization of layered coding systems. This is further discussed in chapter 7.

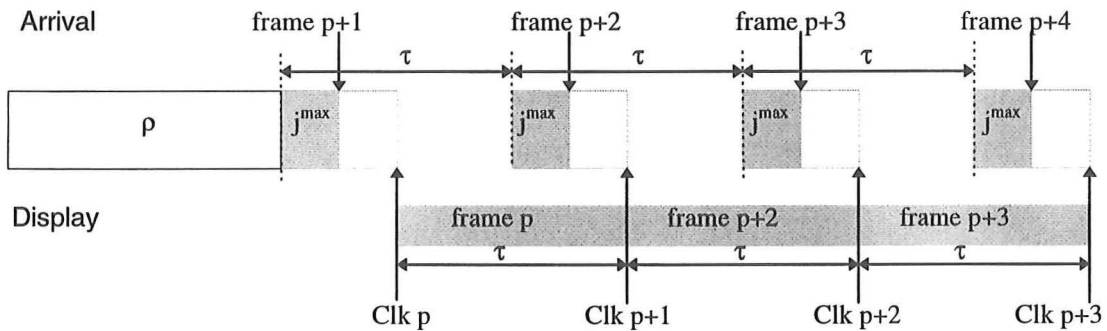


Figure 4.8 : Slipping a media frame with elapsed time $\rho + j^{\max}$

4.5.2 Pause

Similar to a media frame slip, there is an uncertainty region for a media frame pause. The earliest time at which a media frame pause can occur is depicted in figure 4.9. Frame $n+1$ arrives at the latest possible elapsed time. If the time interval η is fractionally higher than

zero, frame $n+1$ has to be paused and frame n should be re-displayed during the interval between $\text{Clk } n+1$ and $\text{Clk } n+2$.

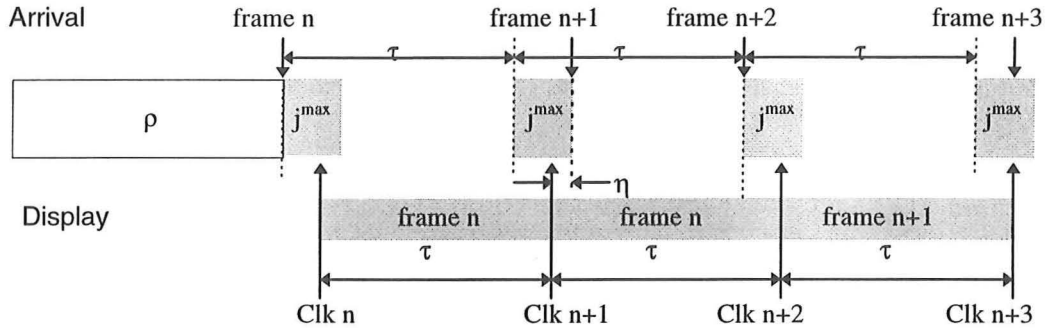


Figure 4.9 : Earliest pause ($\eta > 0, \eta \approx 0$)

Figure 4.10 depicts an intermediate pause. In this situation the time interval η can be any value between 0 and j^{\max} . Media frame $m+1$ which arrives after the occurrence of $\text{Clk } m+1$ is paused.

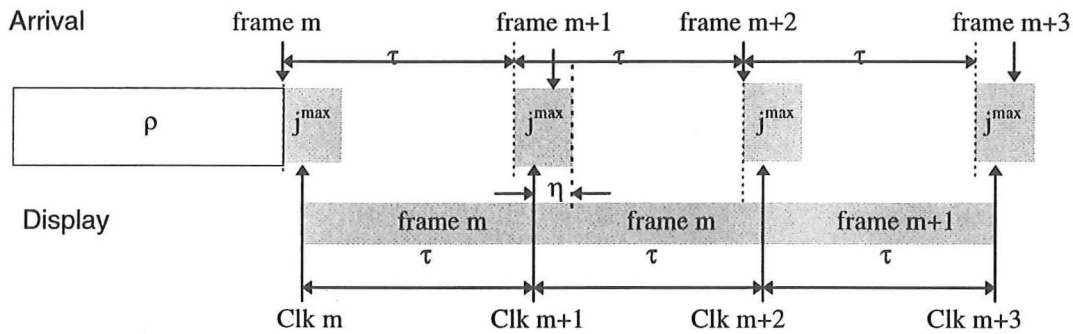


Figure 4.10 : Intermediate Pause ($0 < \eta < j^{\max}$)

However, if frames continuously arrive at the earliest possible elapsed time, a media frame pause will not occur until the situation depicted in figure 4.11. The time interval η should be fractionally above j^{\max} . Frame $p+1$ which arrives later than the occurrence of $\text{Clk } p+1$ is paused and frame p is re-displayed between the interval $\text{Clk } p+1$ and $\text{Clk } p+2$.

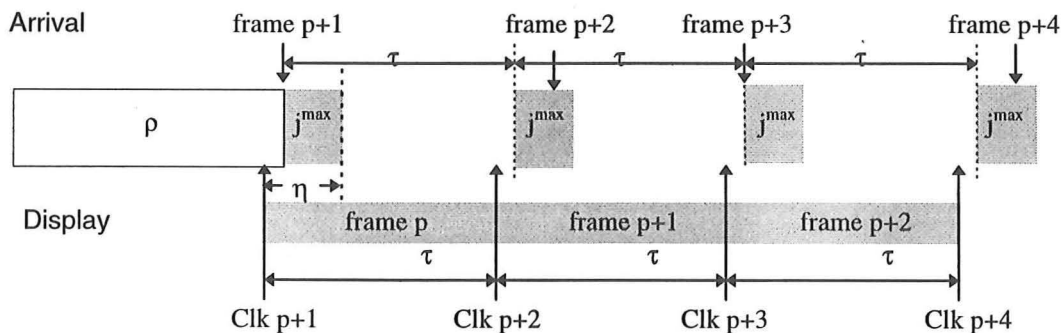


Figure 4.11 : Latest pause ($\eta > j^{\max}, \eta \approx j^{\max}$)

4.6 Slip /pause Frequency

Figure 4.12 depicts delays and jitters of some common media streams. In the figure, the unshaded regions represent the static components of the end-to-end delay and the shaded regions represent the variable components (jitter) of the end-to-end delay similar to figure 4.5.

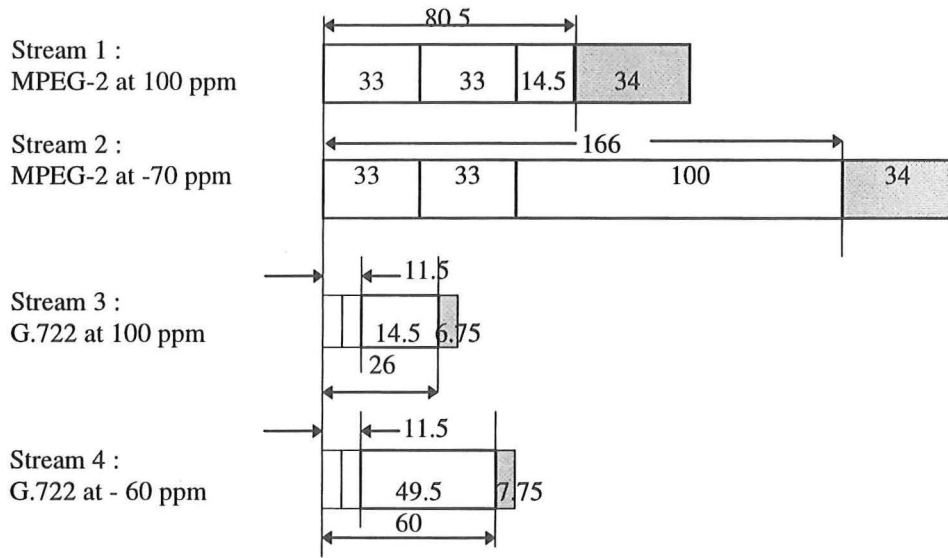


Figure 4.12 : Delays and jitters of video and audio streams

Figure 4.13 plots the end-to-end delay of the media streams as a function of real time. The minimum and maximum end-to-end delays of streams in figure 4.13 are as same as figure 4.12. In addition, in figure 4.13, the end-to-end delay variation with time is also depicted. In the MPEG-2 standard, a sender to receiver clock frequency deviation of 100 parts per million (ppm) is regarded as a typical maximum value. If the sender to receiver clock frequency deviation is 100 ppm, the receiver clock would drift by 100 seconds every million seconds and it would drift by τ (33 ms) in 330 seconds. Thus, a media frame slip or pause could occur approximately every 10,000 frames.

Assume that, the first frame of media stream 1 is displayed at an end-to-end delay of 80.5 ms. As the sender to receiver clock deviation is positive, the receiver clock lags the sender clock and the end-to-end delay of the subsequent frames keep on progressively increasing in proportion to the slip delay. After 335 seconds, the slip delay of the media frame becomes equal to $\tau + j^{\max}$ (33.5 ms), and the process enters the uncertainty region. This uncertainty region will prevail until the clock drifts further by another 0.5 ms. The time period of the uncertainty region is 5s, for clock asynchrony of 100 ppm. Somewhere in this region, a

media frame will have to be slipped. Once a media frame is slipped the end-to-end delay of the media frame drops back to 80.5 ms.

As the sender to receiver clock frequency deviation is negative, the receiver clock runs faster than the sender clock for stream 2. Therefore, the end-to-end delay keeps on reducing progressively as shown. At around 475 seconds, a media frame has to be paused as a frame may not arrive in time for decoding. The media frame pause increases the end-to-end delay of all the subsequent frames. The audio streams also behave similarly but end-to-end delays and jitters are different.

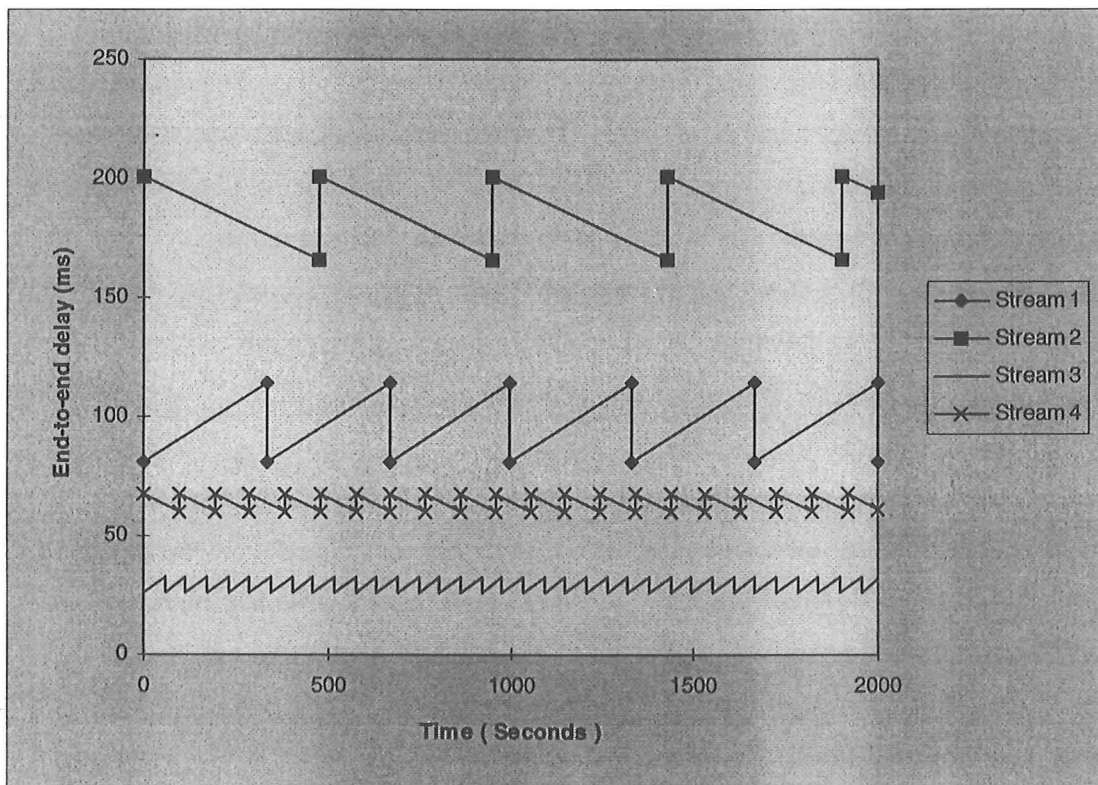


Figure 4.13 : Slipping and pausing intervals of media streams depicted in figure 4.12 at different sender and receiver clock asynchrony

4.7 Inter-media Synchronization

Ideally, two streams produced by a single source such as audio and video should be perfectly synchronized during playout. A number of proposals for inter-media synchronization aim to achieve perfect synchronization [Anderson *et. al.*, 1991]. However, it is impossible to achieve perfect inter-media synchronization between two media streams as there is an uncertainty or jitter in the end-to-end delay (J^{\max}) of any media stream.

Furthermore, in section 2.5.3, it was shown that video can lead / lag audio by

90 ms / 60 ms with only a detected error of 20%. Thus at playout, a stream may have a maximum allowable lead as well as a maximum allowable lag with respect to another stream without unacceptable quality degradation. These allowable skew variations are called skew tolerances. The following notation is used to denote skew tolerances.

$S_{2,1}$ = Maximum allowable lead of stream 2 with respect to stream 1

$S_{1,2}$ = Maximum allowable lead of stream 1 with respect to stream 2

Inter-media synchronization is the process of synchronizing media streams within acceptable skew tolerances. If the media streams are not synchronized within the above skew tolerances, it signifies a lower QoS. Media combining, media multiplexing and common routing are mechanisms suggested for ensuring identical transfer delays of media streams [Campbell *et. al.*, 1992; Pingali *et. al.*, 1991]. However, the media streams may still have different encoder and decoder processing delays and as a result the above three mechanisms cannot guarantee inter-media synchronization. In master slave synchronization [Anderson *et. al.*, 1991], slave media frames are slipped or paused according to the master stream. One of the main drawbacks of this method is that, if an interruption occurs to the master stream, all the slave streams are also affected.

Little *et. al.* (1992) had previously defined target skew tolerance values which were symmetric. These skew tolerance values did not define the constrained skew, but rather the target skew between different streams, which their synchronization mechanism then tried to maintain on a best effort basis. However, there are no known multimedia services that require the playout of different media to be intentionally skewed. In all cases, the target skew should be zero.

Ravindran *et. al.* (1993) proposed a synchronization mechanism, where the clocks of the sender and receiver were synchronized, and a timestamp was appended to each data unit at the sender, to indicate when it should be played out at the receiver. They defined a *maximum skew parameter* which represented the difference between the timestamp's value and the value on the receiver's clock when a data unit was presented. The maximum skew parameters of different streams formed a vector, which they referred to as the *divergence vector*.

However, skew tolerances exist between each possible pair of streams in a session, and this cannot be properly expressed by a divergence vector. Indeed, humans perceive the skew measured between media streams, rather than with respect to a time base, so that to tightly constrain the skew between two media, they both must be tightly constrained to the time base. Arguably, the divergence vector does not directly model the attributes of synchronization that perceived by the user, and is inappropriate for expressing synchronization QoS. Also, if a priori knowledge about the characteristics of the media stream such as delay and jitter is available, the overhead of sending time stamps could be eliminated [Naish, 1994a].

4.8 Static Delay Compensation (SDC)

We propose the mechanism, Static Delay Compensation (SDC) to achieve inter-media synchronization of media streams, within given skew tolerances. In SDC, by adding a static delay to the stream with lower end-to-end delay, two media streams can be synchronized [Naish, *et. al.*, 1995]. Once the bearer connections are established after adding the necessary static delays, the media streams are handled independently with no more synchronization processing requirements. Advantages of SDC over other synchronization mechanisms are that, firstly, there is no exchange of time stamps between the sender and the receiver, and secondly, interruption of a stream has no effect on other streams. Chapter 5 and chapter 6 discuss how SDC can be applied to synchronize two and three media streams, respectively. Chapter 7 elaborates on how synchronization of multiple media streams as well as layered coding systems can be achieved by applying SDC.

CHAPTER 5

STATIC DELAY COMPENSATION ; A MECHANISM FOR SYNCHRONIZING TWO MEDIA STREAMS

In Static Delay Compensation (SDC), inter-media synchronization of two media streams within acceptable skew tolerance is achieved by adding a static delay to the stream with lower end-to-end delay. Inter-media synchronization does not require perfect synchrony between the two streams. Instead a stream may be allowed to lead or lag the other stream within acceptable skew. Thus, two skew constraints representing lead and lag of a stream with respect to the other need to be satisfied when synchronizing two streams. In this chapter, four cases that need to be considered in synchronizing two streams are discussed and examples of the application of Static Delay Compensation are presented.

5.1 Static Delay Compensation (SDC)

We propose the mechanism, Static Delay Compensation (SDC) to achieve inter-media synchronization of media streams. In SDC, by adding a static delay to the stream with lower end-to-end delay, two media streams can be synchronized. In synchronizing two media streams the following two skew constraints need to be satisfied.

Skew Constraint 1 : Maximum lead of stream 2 with respect to stream 1 should be less than the corresponding skew tolerance value.

Skew Constraint 2 : Maximum lead of stream 1 with respect to stream 2 should be less than the corresponding skew tolerance value.

Considering the above skew constraints, there are four possible cases when synchronizing two streams, namely

- Case I : Both skew constraints satisfied
- Case II : One skew constraint satisfied and the streams are able to be synchronized
- Case III : One skew constraint satisfied and the streams are not able to be synchronized
- Case IV : Both skew constraints are not satisfied

5.2 Case I : Both Skew Constraints Satisfied

If both skew constraints are satisfied then the streams are synchronized already. The following discussion depicts a situation where both skew constraints are satisfied. Consider the two media streams represented in figure 5.1. The following notation is used in the next few sections. The symbols ρ_i and σ_i , $i = 1, 2, 3, \dots, n$ represent the minimum and maximum end-to-end delays of the original streams and symbols D^{\min} and D^{\max} represent the minimum and maximum end-to-end delays of the synchronized streams. If the streams are already synchronized then no static delay is added to each of the streams and the minimum and maximum end-to-end delays of the original streams are equivalent to the minimum and maximum end-to-end delays of the synchronized streams. Hence $D^{\min}_1 = \rho_1$, $D^{\min}_2 = \rho_2$, $D^{\max}_1 = \sigma_1$ and $D^{\max}_2 = \sigma_2$

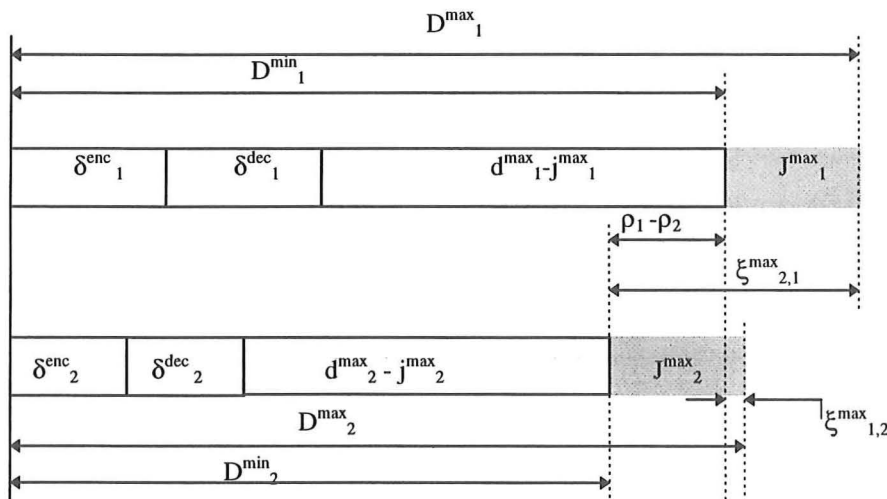


Figure 5.1 : Delays and jitters of stream 1 and stream 2

Maximum lead of stream 2 with respect to stream 1 = $D^{\max}_1 - D^{\min}_2 = \xi^{\max}_{2,1}$

Maximum allowable lead of stream 2 with respect to stream 1 (corresponding skew tolerance value) = $S_{2,1}$

Therefore skew constraint 1 can be stated as,

$$\xi_{2,1}^{\max} \leq S_{2,1} \quad (5.1)$$

Synchronization coefficient $\gamma_{2,1}$ is defined as the extent to which the first skew constraint has been exceeded.

$$\gamma_{2,1} = \xi_{2,1}^{\max} - S_{2,1} \quad (5.2)$$

If $\gamma_{2,1}$ is positive, the first skew constraint has been exceeded by $\gamma_{2,1}$ while if $\gamma_{2,1}$ is negative then the first skew constraint is satisfied with a margin of $-\gamma_{2,1}$.

From the figure, $\xi_{2,1}^{\max} = \rho_1 - \rho_2 + J_1^{\max}$ and therefore,

$$\gamma_{2,1} = \rho_1 - \rho_2 + J_1^{\max} - S_{2,1} \quad (5.3)$$

Maximum lead of stream 1 with respect to stream 2 = $D_2^{\max} - D_1^{\min} = \xi_{1,2}^{\max}$

Maximum allowable lead of stream 1 with respect to stream 2 (corresponding skew tolerance value) = $S_{1,2}$

Therefore skew constraint 2 can be stated as,

$$\xi_{1,2}^{\max} \leq S_{1,2} \quad (5.4)$$

The synchronization coefficient $\gamma_{1,2}$ is defined as the extent to which the second skew constraint has been exceeded.

$$\gamma_{1,2} = \xi_{1,2}^{\max} - S_{1,2} \quad (5.5)$$

If $\gamma_{1,2}$ is positive the second skew constraint has been exceeded by $\gamma_{1,2}$ while if $\gamma_{1,2}$ is negative then the second skew constraint is satisfied with a margin of $-\gamma_{1,2}$.

From the figure, $\xi_{1,2}^{\max} = \rho_2 - \rho_1 + J_2^{\max}$ and therefore

$$\gamma_{1,2} = \rho_2 - \rho_1 + J_2^{\max} - S_{1,2} \quad (5.6)$$

If the streams are already synchronized then both skew constraints should be satisfied already. If the first skew constraint is satisfied then by definition, $\gamma_{2,1} \leq 0$ and if the second skew constraint is satisfied then, by definition, $\gamma_{1,2} \leq 0$. **Therefore, if $\gamma_{2,1} \leq 0$ and $\gamma_{1,2} \leq 0$, the streams are already synchronized.**

5.3 Case II : One Skew Constraint Satisfied and the Streams are able to be Synchronized

If one skew constraint is satisfied while the other is not, either $\gamma_{1,2}$ or $\gamma_{2,1}$ is positive while the other is negative. Assume that the first skew constraint is not satisfied and the second skew constraint is satisfied. Then $\gamma_{2,1} > 0$ and $\gamma_{1,2} \leq 0$. Here synchronization is attempted by adding a static delay $\beta_{2,1}^{\min}$ to stream 2, as shown in figure 5.2 below.

It should be noted that there is a range of possible static delay values that are applicable with the given skew tolerances. The minimum static delay is represented by $\beta_{2,1}^{\min}$. In figure 5.2, by adding a static delay $\beta_{2,1}^{\min}$ to stream 2, the maximum lead of stream 2 with respect to stream 1 is made equal to $S_{2,1}$. Note that no static delay is added to stream 1 and therefore, $D_1^{\min} = \rho_1$ and $D_1^{\max} = \sigma_1$

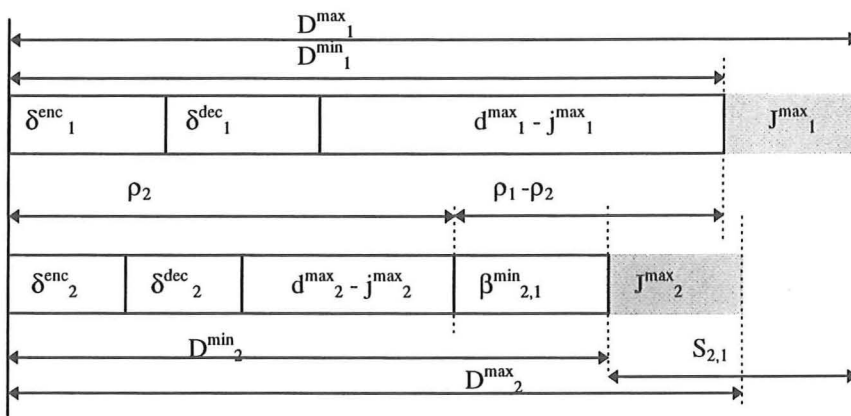


Figure 5.2 : Minimum static delay to be added to stream 2 to synchronize it with stream 1

In figure 5.2,

$$\beta_{2,1}^{\min} = \rho_1 - \rho_2 + J_1^{\max} - S_{2,1} \quad (5.7)$$

By the definition of $\gamma_{2,1}$ given in equation 5.3 and equation 5.7,

$$\beta_{2,1}^{\min} = \gamma_{2,1} \quad (5.8)$$

Synchronization coefficients of the original streams are denoted by $(\gamma_{2,1})_0$ and $(\gamma_{1,2})_0$ where the subscript 0 signifies zeroth iteration. Thus by definition,

$$(\gamma_{2,1})_0 = \rho_1 - \rho_2 + J^{\max}_1 - S_{2,1} \quad (5.9)$$

$$(\gamma_{1,2})_0 = \rho_2 - \rho_1 + J^{\max}_2 - S_{1,2} \quad (5.10)$$

After adding the static delay, the synchronization coefficients are given by

$$(\gamma_{2,1})_1 = \rho_1 - (\rho_2 + \beta^{\min}_{2,1}) + J^{\max}_1 - S_{2,1} \quad (5.11)$$

$$(\gamma_{1,2})_1 = (\rho_2 + \beta^{\min}_{2,1}) - \rho_1 + J^{\max}_2 - S_{1,2} \quad (5.12)$$

where subscript 1 denotes the first iteration.

By substituting for $\beta^{\min}_{2,1}$ from equation 5.8 and rearranging the above two equations we have ,

$$(\gamma_{2,1})_1 = \rho_1 - \rho_2 + J^{\max}_1 - S_{2,1} - (\gamma_{2,1})_0 \quad (5.13)$$

$$(\gamma_{1,2})_1 = \rho_2 - \rho_1 + J^{\max}_2 - S_{1,2} + (\gamma_{2,1})_0 \quad (5.14)$$

By substituting from equations 5.9 and 5.10 we have

$$(\gamma_{2,1})_1 = (\gamma_{2,1})_0 - (\gamma_{2,1})_0 = 0 \quad (5.15)$$

$$(\gamma_{1,2})_1 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0 \quad (5.16)$$

Therefore as $(\gamma_{2,1})_1 = 0$, the first skew constraint is satisfied. If $(\gamma_{1,2})_0 + (\gamma_{2,1})_0 \leq 0$ then second skew constraint is still satisfied after the first iteration. Hence the media streams can be synchronized by adding a static delay β_2 equivalent to $\gamma_{2,1}$ to stream 2, if $\gamma_{2,1} \geq 0$, $\gamma_{1,2} \leq 0$, and $\gamma_{1,2} + \gamma_{2,1} \leq 0$, where all the synchronization coefficients refer to the zeroth iteration. In other words, if $\gamma_{2,1} \geq 0$, $\gamma_{1,2} \leq 0$, and $|\gamma_{2,1}| \leq |\gamma_{1,2}|$, media streams can be synchronized by adding a static delay. This case also applies to $\gamma_{2,1} \leq 0$, $\gamma_{1,2} \geq 0$, and $|\gamma_{1,2}| \leq |\gamma_{2,1}|$.

5.4 Case III : One Skew Constraint Satisfied and the Streams are not able to be Synchronized

If $(\gamma_{2,1})_0 + (\gamma_{2,1})_0 > 0$ then second skew constraint is not satisfied after the first iteration, as synchronization coefficient $(\gamma_{1,2})_1 > 0$. As $(\gamma_{2,1})_1 = 0$ the first skew constraint is satisfied. Now a

static delay $(\beta^{\min}_{1,2})_1$ equivalent to $(\gamma_{1,2})_1$ or $(\gamma_{1,2})_0 + (\gamma_{2,1})_0$ should be added to stream 1 to synchronize the streams. Thus

$$(\beta^{\min}_{1,2})_1 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0 \quad (5.17)$$

After adding this static delay synchronization coefficients are given by

$$(\gamma_{2,1})_2 = (\rho_1 + (\beta^{\min}_{1,2})_1) - (\rho_2 + \beta^{\min}_{2,1}) + J^{\max}_1 - S_{2,1} \quad (5.18)$$

$$(\gamma_{1,2})_2 = (\rho_2 + \beta^{\min}_{2,1}) - (\rho_1 + (\beta^{\min}_{1,2})_1) + J^{\max}_2 - S_{1,2} \quad (5.19)$$

By substituting for $\beta^{\min}_{2,1}$ and $(\beta^{\min}_{1,2})_1$ from equations 5.8 and 5.17, and for $(\gamma_{1,2})_0$ and $(\gamma_{2,1})_0$ from equations 5.9 and 5.10, we have,

$$(\gamma_{2,1})_2 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0 \quad (5.20)$$

$$(\gamma_{1,2})_2 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0 - \{(\gamma_{1,2})_0 + (\gamma_{2,1})_0\} = 0 \quad (5.21)$$

At the second iteration, as $(\gamma_{1,2})_0 + (\gamma_{2,1})_0 > 0$, the first skew constraint will not be satisfied. Thus it is clearly seen that both skew constraints cannot be satisfied as $\gamma_{1,2}$ and $\gamma_{2,1}$ will alternate at each iteration. Hence media streams are not able to be synchronized if $\gamma_{2,1} \geq 0$, $\gamma_{1,2} \leq 0$, $\gamma_{1,2} + \gamma_{2,1} > 0$, **where all the synchronization coefficients refer to the zeroth iteration.** In other words, if $\gamma_{2,1} \geq 0$, $\gamma_{1,2} \leq 0$ and $|\gamma_{2,1}| \geq |\gamma_{1,2}|$ media streams cannot be synchronized. This case also applies to $\gamma_{2,1} \leq 0$, $\gamma_{1,2} \geq 0$ and $|\gamma_{1,2}| \geq |\gamma_{2,1}|$.

5.5 Case IV : Both Skew Constraints not Satisfied

If both skew constraints are not satisfied then both $(\gamma_{1,2})_0$ and $(\gamma_{2,1})_0$ are positive. A static delay equivalent to $(\gamma_{2,1})_0$ should be added to stream 2 to satisfy the first skew constraint. If the above mentioned static delay is added, the synchronization coefficients at the first iteration will be equal to $(\gamma_{2,1})_1 = 0$ and $(\gamma_{1,2})_1 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0$. As $(\gamma_{1,2})_0 + (\gamma_{2,1})_0 > 0$, similar to an argument in the previous section it can be shown that the streams are not able to be synchronized.

Therefore if $\gamma_{1,2} > 0$ and $\gamma_{2,1} > 0$, **the streams cannot be synchronized**, where the synchronization coefficients refer to the zeroth iteration. This result can be seen intuitively, as

opposing static delays need to be added to each stream. The total jitter of the end-to-end delay of any media stream between any two points across the globe is low, if fibre optics is being used as the transmission medium and the number of ATM switching stages is not large. Thus, the probability of case IV occurring is extremely low if the above underlying technologies are used. This will be shown in the examples given in section 5.7.

5.6 Summary of All Possible Situations

The figure 5.3 summarizes all possible cases.

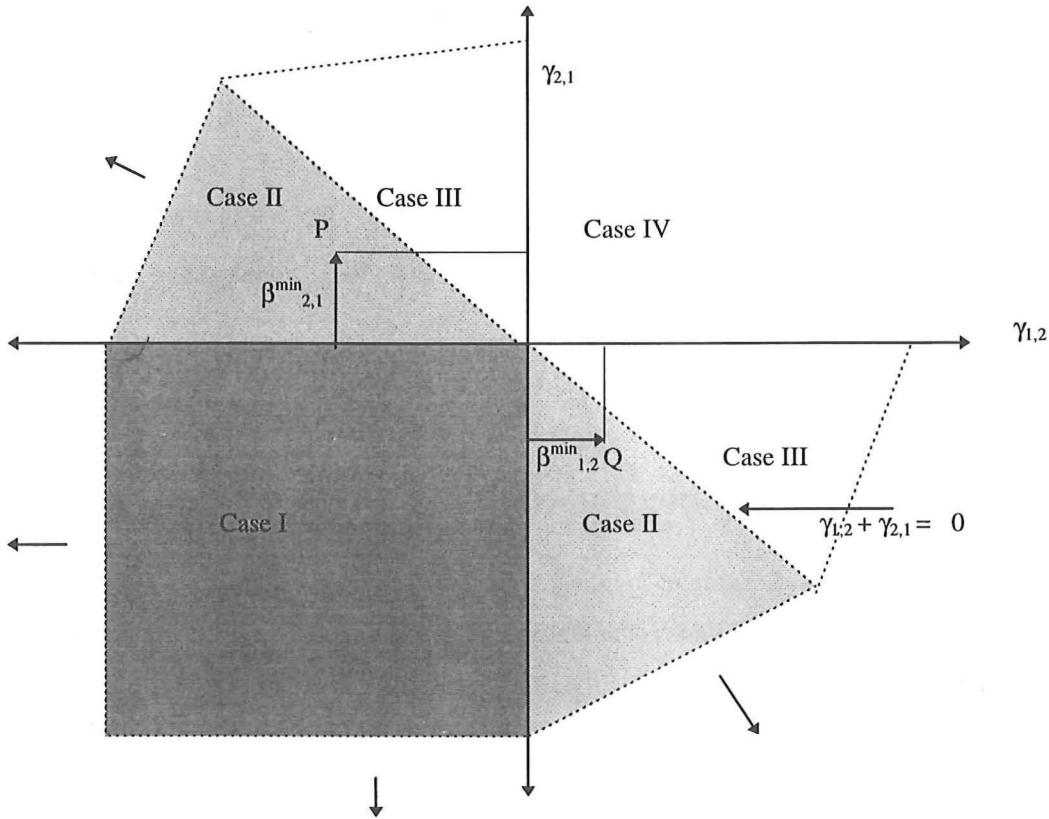


Figure 5.3 : Summary of all possible cases for synchronization

In case I, both skew constraints are satisfied, $\gamma_{2,1} \leq 0$ and $\gamma_{1,2} \leq 0$ and the streams are already synchronized. This case is represented by the rectangular shaded region.

In case II, one skew constraint is satisfied and the streams are able to be synchronized. This case is represented by triangular shaded regions. In the situation shown by point P, $\gamma_{2,1} > 0$, $\gamma_{1,2} < 0$ and $\gamma_{1,2} + \gamma_{2,1} < 0$ and therefore $\beta_{2,1}^{min} = \gamma_{2,1}$. In the situation shown by point Q, $\gamma_{2,1} < 0$, $\gamma_{1,2} > 0$, $\gamma_{1,2} + \gamma_{2,1} < 0$ and $\beta_{1,2}^{min} = \gamma_{1,2}$.

In case III, one skew constraint is satisfied and the streams are not able to be synchronized. This case is represented by triangular unshaded regions.

In case IV, both skew constraints are not satisfied, and the streams cannot be synchronized. This case is represented by the rectangular unshaded region.

In summary, if $\gamma_{1,2} + \gamma_{2,1} \leq 0$ then media streams are able to be synchronized and if $\gamma_{1,2} + \gamma_{2,1} > 0$ media streams are not able to be synchronized. The static delays that should be added to the two streams can be written as

$$\beta_1 = a_{1,2} * \gamma_{1,2} \text{ and}$$

$$\beta_2 = a_{2,1} * \gamma_{2,1}$$

$$\text{where } a_{1,2} = \begin{cases} 0 & \text{when } \gamma_{1,2} \leq 0 \\ 1 & \text{when } \gamma_{1,2} > 0 \end{cases}$$

$$\text{and } a_{2,1} = \begin{cases} 0 & \text{when } \gamma_{2,1} \leq 0 \\ 1 & \text{when } \gamma_{2,1} > 0 \end{cases}$$

5.7 Examples of Application of Static Delay Compensation

In this section three numerical examples are discussed. Consider a multimedia conference which is set up between Christchurch and Sydney. Each bearer connection of the multimedia conference could be set up either via satellite or via the fibre optic network. In this example only a one way (simplex) call is considered for simplicity.

5.7.1 Example 1

Say a MPEG-1 video stream as well as an audio stream are transmitted as shown in figure 5.4, and the audio stream has to be converted from hi-fidelity to G.722. The encoder and the decoder processing delays of the video stream are equivalent to 33 ms each [Jeffay *et. al.*, 1995]. The encoder and decoder processing delays of the audio stream are taken as 33 ms and 5.75 ms respectively [Jeffay *et. al.*, 1995]. The code converter delay of the audio stream is taken as 25 ms, which assumes that the conversion process is done without fully decoding the original signal. As the fibre optic length is about 2900 km the transfer delay of both streams would be equivalent to $2900 * 5 / 1000 \text{ ms} = 14.5 \text{ ms}$, at $5 \mu\text{s/km}$ transfer delay.

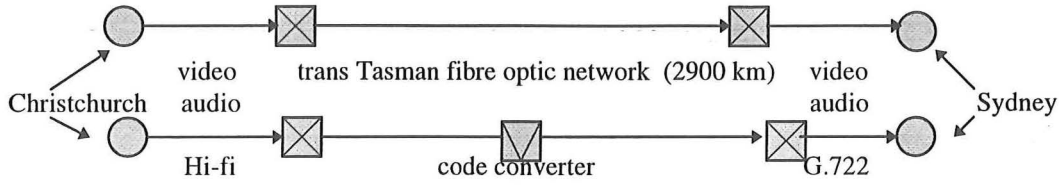


Figure 5.4 : Both audio and video streams transmitted via a fibre optic network (case I)

The delays and jitters of the two streams are indicated in figure 5.5. The jitter in the end-to-end delay of the audio stream is higher because there are two slip/pause stages. The jitter in the end-to-end delay of the two streams are given by

$$J_{\text{audio}}^{\text{max}} = 33 \text{ ms} + 500 \mu\text{s} \times 2 = 34 \text{ ms}$$

$$J_{\text{audio}}^{\text{max}} = 33 \text{ ms} + 250 \mu\text{s} \times 2 + 5.75 \text{ ms} + 250 \mu\text{s} \times 2 = 39.75 \text{ ms}$$

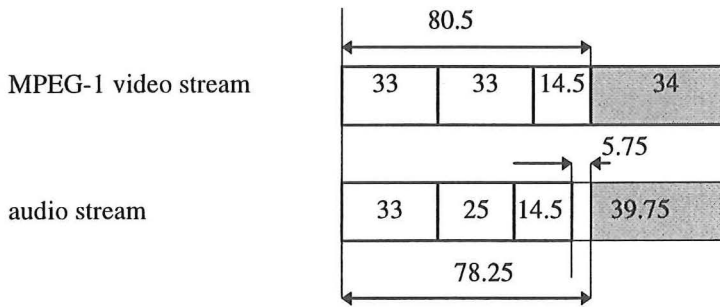


Figure 5.5 : Delays and jitters of video and audio streams depicted in figure 5.4

Without unacceptable quality degradation the video stream is allowed to lead the audio stream by 90 ms and the audio stream is allowed to lead the video stream by 60 ms [Steinmetz *et. al.*, 1993] . Hence skew tolerance values are taken as $S_{1,2} = 90 \text{ ms}$ and $S_{2,1} = 60 \text{ ms}$. By the definition of synchronization coefficients

$$\gamma_{1,2} = -52.5 \quad (5.22)$$

$$\gamma_{2,1} = -39.75 \quad (5.23)$$

Since both synchronization coefficients are negative the streams are already synchronized.

5.7.2 Example 2

Now consider a situation where a G.722 audio stream is transmitted without any code conversion as shown in figure 5.6. The delays and jitters of the two streams can be represented as shown in

figure 5.7. The encoder and decoder delays of G. 722 audio stream are taken as 5.75 ms each. The J^{\max} term of G.722 audio stream is equal to 6.75 ms as shown in section 4.4.

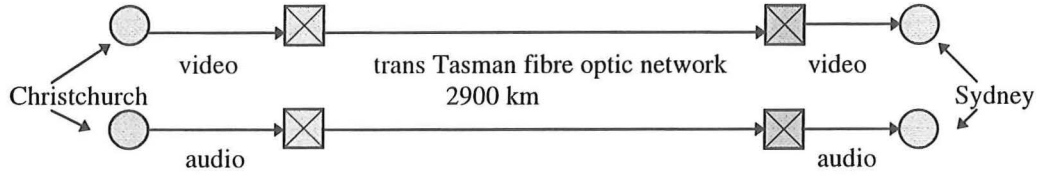


Figure 5.6 : Both audio and video streams transmitted via a fibre optic link (case II)

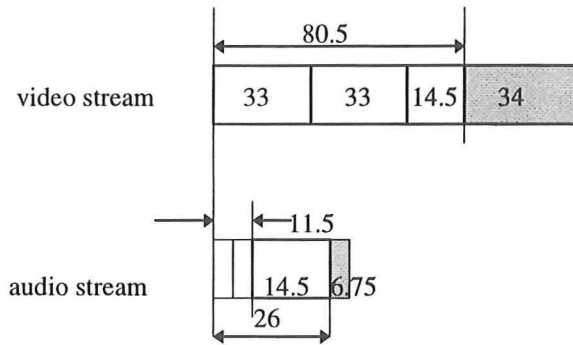


Figure 5.7 : Video and audio streams for situation depicted in figure 5.6

Now the synchronization coefficients are equivalent to

$$\gamma_{1,2} = -137.75 \quad (5.24)$$

$$\gamma_{2,1} = 28.5 \quad (5.25)$$

$\gamma_{1,2} + \gamma_{2,1} = -109.25$, and hence the streams can be synchronized by addition a static delay of 28.5 ms to the audio stream.

5.7.3 Example 3

Now consider the situation depicted in figure 5.8, where a MPEG-1 video stream and a G.722 audio stream are transmitted via satellite. The satellite hop incurs a maximum delay of 270 ms and a maximum delay-jitter of 40 ms. Thus the minimum transfer delay is equivalent to 230 ms. Figure 5.9 represents the components of delays and jitters of the video and audio streams.

Now the synchronization coefficients are equivalent to,

$$\gamma_{1,2} = -57.75 \quad (5.26)$$

$$\gamma_{2,1} = 108.5 \quad (5.27)$$

$\gamma_{1,2} + \gamma_{2,1} = 50.75$, and hence the streams cannot be synchronized. Therefore when satellite communications are used to transmit both audio and video streams, synchronization cannot be achieved within acceptable skew tolerances.

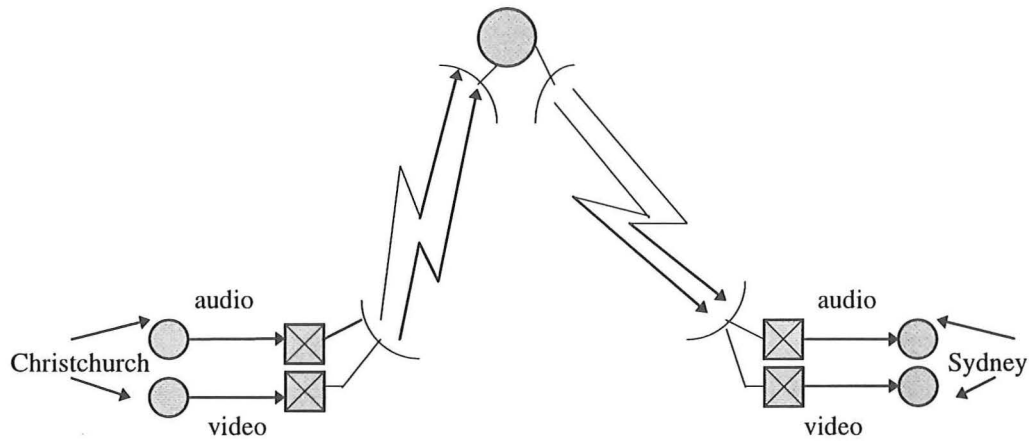


Figure 5.8 : Both audio and video streams are transmitted via a satellite link

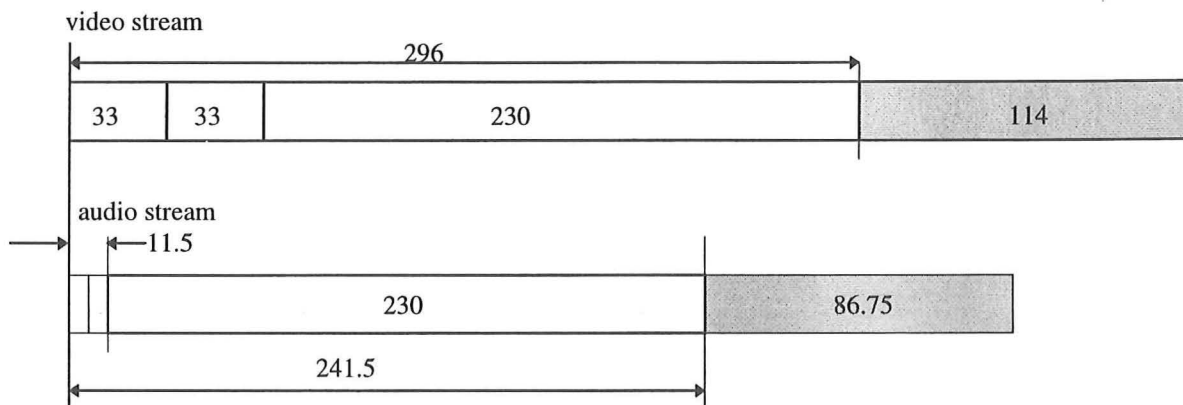


Figure 5.9 : Components of end-to-end delay of audio and video streams

CHAPTER 6

STATIC DELAY COMPENSATION ; A MECHANISM FOR SYNCHRONIZING THREE MEDIA STREAMS

In this chapter, a mechanism is described to synchronize three media streams in broadband communications, which extends the mechanism developed for synchronization of two media streams. In synchronizing three media streams skew constraints have to be preserved between all combinations of pairs of streams. In Static Delay Compensation (SDC), inter-media synchronization within acceptable skew tolerances is achieved by adding static delays to the streams with lower end-to-end delays. Finally, examples are given for a representative set of cases of synchronization coefficients and an equation is derived to calculate the static delay to be added to each media stream, in synchronizing three media streams.

6.1 Synchronization of Three or more Media Streams

In an application where multiple audio and video streams are transmitted all media may have to be played out in synchrony. Also in other specialized applications either multiple streams originating from different sources may have to be played in synchrony at a single destination or a single stream may have to be played in synchrony in multiple destinations [Escobar *et. al.*, 1992].

The layers of a layered coding system also have to be played out synchronously at the receiver.

The mechanism developed in the previous chapter, provides a foundation for optimally synchronizing multiple media streams. This chapter discusses synchronization of three media streams while synchronization of more than three media streams and synchronization of layered coding systems are discussed in the next chapter.

6.2 Constraints to be Satisfied When Synchronizing Three Media Streams

In this section, the constraints to be satisfied when synchronizing three media streams by the application of Static Delay Compensation is analyzed. Figure 6.1 depicts the delays and the jitters of the three media streams to be synchronized.

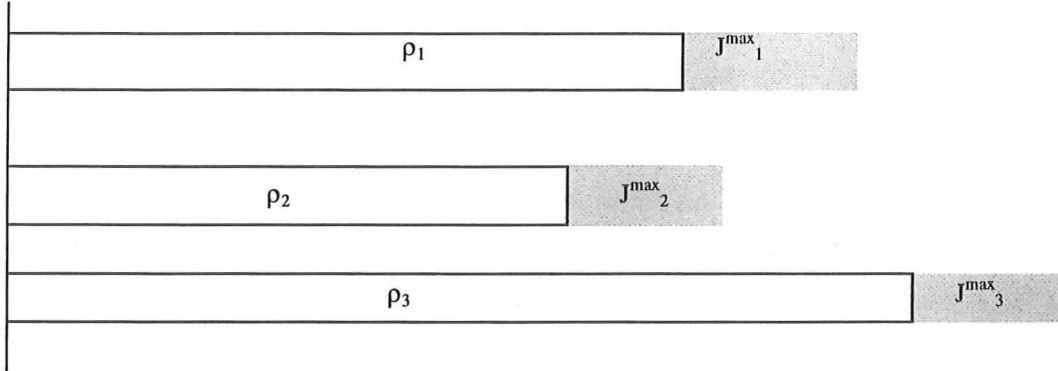


Figure 6.1: Delays and jitters of the three streams

By the definition of synchronization coefficients, from section 5.2, for synchronization of stream 1 and 2

$$(\gamma_{1,2})_0 = \rho_2 - \rho_1 + J_2^{\max} - S_{1,2} \quad (6.1)$$

$$(\gamma_{2,1})_0 = \rho_1 - \rho_2 + J_1^{\max} - S_{2,1} \quad (6.2)$$

For synchronization of stream 1 and 3

$$(\gamma_{1,3})_0 = \rho_3 - \rho_1 + J_3^{\max} - S_{1,3} \quad (6.3)$$

$$(\gamma_{3,1})_0 = \rho_1 - \rho_3 + J_1^{\max} - S_{3,1} \quad (6.4)$$

Similarly for synchronization of stream 2 and 3

$$(\gamma_{2,3})_0 = \rho_3 - \rho_2 + J_3^{\max} - S_{2,3} \quad (6.5)$$

$$(\gamma_{3,2})_0 = \rho_2 - \rho_3 + J_2^{\max} - S_{3,2} \quad (6.6)$$

Synchronization is possible between the pair of stream 1 and 2 only if $(\gamma_{1,2})_0 + (\gamma_{2,1})_0 \leq 0$.

Similarly synchronization is possible between the pair of stream 1 and 3 and between the pair of streams 2 and 3 only if $(\gamma_{1,3})_0 + (\gamma_{3,1})_0 \leq 0$ and $(\gamma_{2,3})_0 + (\gamma_{3,2})_0 \leq 0$, respectively.

Consider the cartesian diagram shown in figure 6.2, which is used to depict the synchronization coefficients. The x axis is simultaneously used to mark $\gamma_{1,2}, \gamma_{2,3}, \gamma_{3,1}$, while the y axis is simultaneously used to mark $\gamma_{2,1}, \gamma_{3,2}$ and $\gamma_{1,3}$. The synchronization point $P_{1,2}$ denotes the coordinates $(\gamma_{1,2}, \gamma_{2,1})$. The synchronization points $P_{2,3}$ and $P_{3,1}$ denote the coordinates $(\gamma_{2,3}, \gamma_{3,2})$ and $(\gamma_{3,1}, \gamma_{1,3})$, respectively.

As described in section 5.6, if a synchronization point lies in the third quadrant, then the corresponding media streams are already synchronized. If a synchronization point lies in either the second or the fourth quadrant below the $\gamma_{x,y} + \gamma_{y,x} = 0$ line, then the corresponding media streams can be synchronized by adding a static delay to one of the streams. As the synchronization point $(P_{1,2})_0$ lie in the third quadrant, media stream 1 and 2 are already synchronized. As $(P_{3,1})_0$ lie in the third quadrant, media stream 1 and 3 are also synchronized. The synchronization point $(P_{2,3})_0$ lies in the fourth quadrant below the $\gamma_{x,y} + \gamma_{y,x} = 0$ line, and by adding a static delay (β_2) equivalent to $(\gamma_{2,3})_0$ to stream 2, it could be synchronized with stream 3. Figure 6.3 depicts the transformation of synchronization coefficients after adding this static delay.

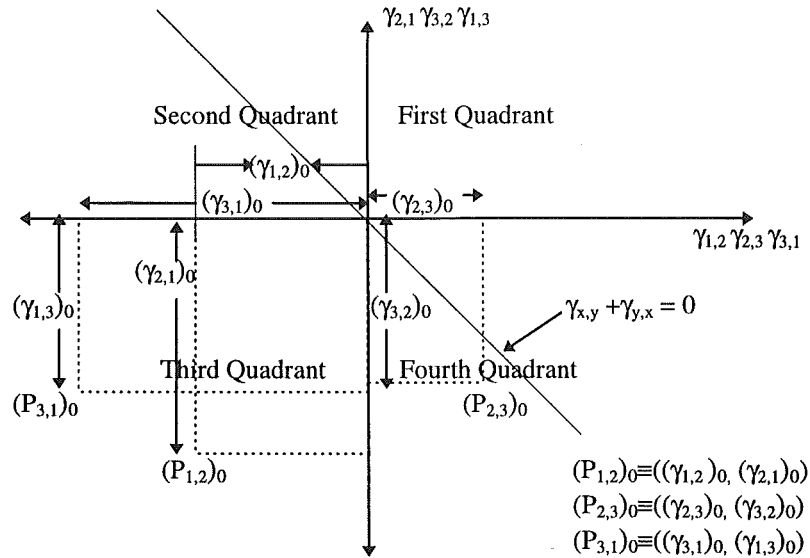


Figure 6.2 : Cartesian diagram showing synchronization coefficients of the three streams

After adding this static delay, the synchronization coefficients $\gamma_{2,3}$ and $\gamma_{3,2}$ are given by

$$(\gamma_{2,3})_1 = \rho_3 - (\rho_2 + \beta_2) + J_{3-}^{\max} - S_{2,3} \quad (6.7)$$

$$(\gamma_{3,2})_1 = (\rho_2 + \beta_2) - \rho_3 + J_{2-}^{\max} - S_{3,2} \quad (6.8)$$

By substituting for β_2 , and rearranging the two equations we have

$$(\gamma_{2,3})_1 = \rho_3 - \rho_2 + J_{3-}^{\max} S_{2,3} - (\gamma_{2,3})_0 \quad (6.9)$$

$$(\gamma_{3,2})_1 = \rho_2 - \rho_3 + J_{2-}^{\max} S_{3,2} + (\gamma_{2,3})_0 \quad (6.10)$$

By substituting from equations 6.5 and 6.6 we have

$$(\gamma_{2,3})_1 = (\gamma_{2,3})_0 - (\gamma_{2,3})_0 = 0 \quad (6.11)$$

$$(\gamma_{3,2})_1 = (\gamma_{3,2})_0 + (\gamma_{2,3})_0 \quad (6.12)$$

The addition of static delay β_2 will also affect the synchronization coefficients $\gamma_{1,2}$ and $\gamma_{2,1}$. The new values of the above two synchronization coefficients are given by

$$(\gamma_{1,2})_1 = (\rho_2 + \beta_2) - \rho_1 + J_{2-}^{\max} S_{1,2} \quad (6.13)$$

$$(\gamma_{2,1})_1 = \rho_1 - (\rho_2 + \beta_2) + J_{1-}^{\max} S_{2,1} \quad (6.14)$$

By substituting for β_2 , and rearranging the two equations we have

$$(\gamma_{1,2})_1 = \rho_2 - \rho_1 + J_{2-}^{\max} S_{1,2} + (\gamma_{2,3})_0 \quad (6.15)$$

$$(\gamma_{2,1})_1 = \rho_1 - \rho_2 + J_{1-}^{\max} S_{2,1} - (\gamma_{2,3})_0 \quad (6.16)$$

By substituting from equations 6.1 and 6.2 we have

$$(\gamma_{1,2})_1 = (\gamma_{1,2})_0 + (\gamma_{2,3})_0 \quad (6.17)$$

$$(\gamma_{2,1})_1 = (\gamma_{2,1})_0 - (\gamma_{2,3})_0 \quad (6.18)$$

In figure 6.3, the synchronization point $(P_{1,2})_1$ lies in the third quadrant and the resultant streams are synchronized. In the example described in this section, the synchronization could be achieved after one iteration. However, in the situation depicted in figure 6.4, the initial position of the synchronization points are such that synchronization cannot be reached after one iteration.

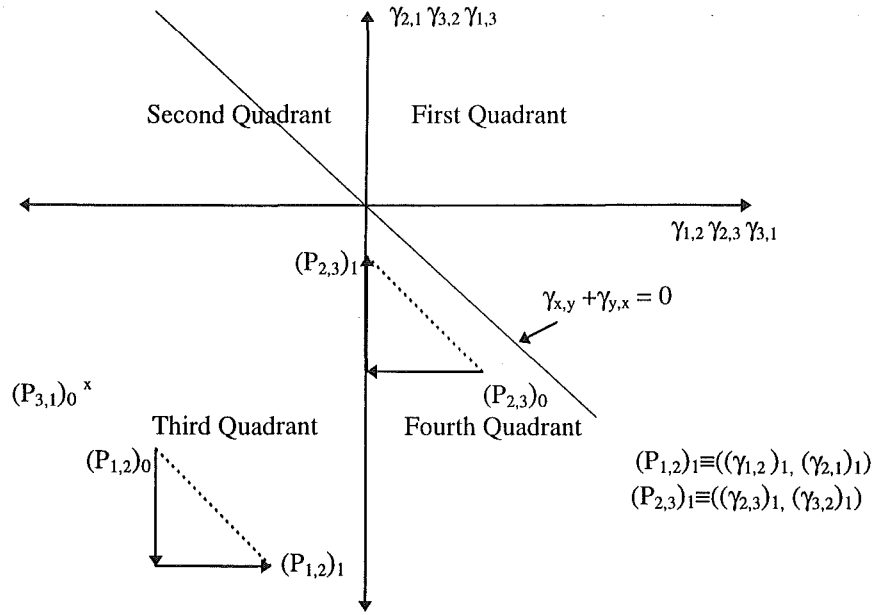


Figure 6.3 : Cartesian diagram showing the transformation of synchronization coefficients when $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 \leq 0$

In figure 6.4, the synchronization point $(P_{1,2})_1$ lies in the fourth quadrant and another static delay (β_1) equivalent to $(\gamma_{1,2})_1$ or $(\gamma_{1,2})_0 + (\gamma_{2,3})_0$ should be added to stream 1 to synchronize stream 1 and 2. After the addition of this static delay the synchronization coefficients are depicted in figure 6.5.

The values of the affected synchronization coefficients, after the second iteration are given by

$$(\gamma_{1,2})_2 = (\rho_2 + \beta_2) - (\rho_1 + \beta_1) + J^{\max}_2 - S_{1,2} \quad (6.19)$$

$$(\gamma_{2,1})_2 = (\rho_1 + \beta_1) - (\rho_2 + \beta_2) + J^{\max}_1 - S_{2,1} \quad (6.20)$$

$$(\gamma_{1,3})_2 = \rho_3 - (\rho_1 + \beta_1) + J^{\max}_3 - S_{1,3} \quad (6.21)$$

$$(\gamma_{3,1})_2 = (\rho_1 + \beta_1) - \rho_3 + J^{\max}_1 - S_{3,1} \quad (6.22)$$

By substituting for β_2 and β_1 and the original synchronization coefficients from equations 6.1, 6.2, 6.3 and 6.4, we have,

$$(\gamma_{1,2})_2 = (\gamma_{1,2})_0 + (\gamma_{2,3})_0 - \{(\gamma_{1,2})_0 + (\gamma_{2,3})_0\} = 0 \quad (6.23)$$

$$(\gamma_{2,1})_2 = (\gamma_{1,2})_0 + (\gamma_{2,1})_0 \quad (6.24)$$

$$(\gamma_{1,3})_2 = (\gamma_{1,3})_0 - (\gamma_{1,2})_0 - (\gamma_{2,3})_0 \quad (6.25)$$

$$(\gamma_{3,1})_2 = (\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 \quad (6.26)$$

The synchronization point $(P_{2,3})_1$ is unaffected during the second iteration and has coefficients $(\gamma_{2,3})_2 = 0$ and $(\gamma_{3,2})_2 = (\gamma_{3,2})_1$

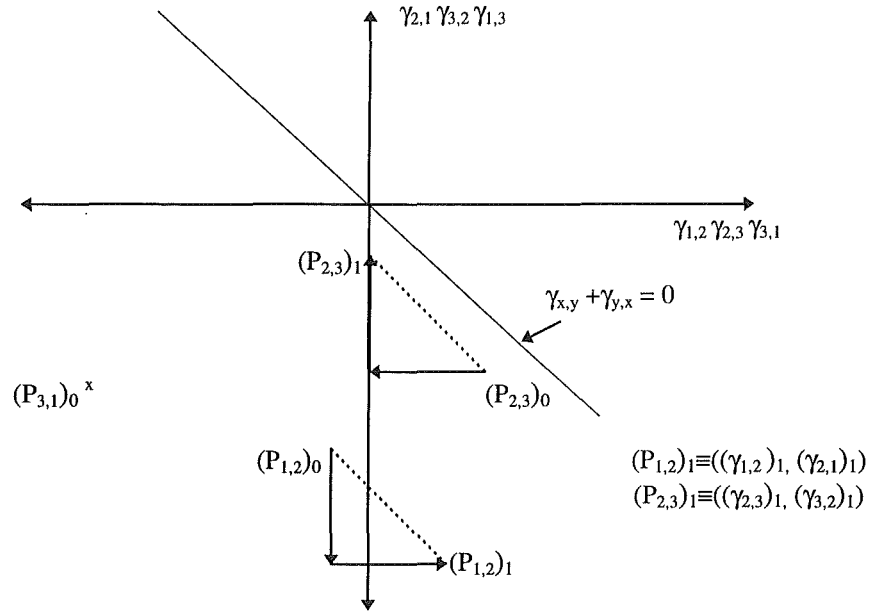


Figure 6.4 : Cartesian diagram showing the transformation of synchronization coefficients when $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 > 0$

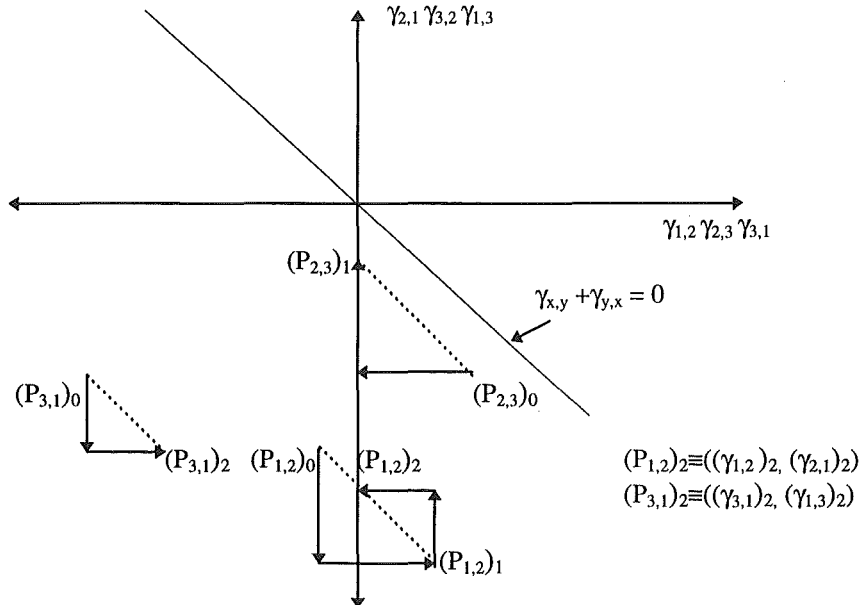


Figure 6.5 : Cartesian diagram showing the second iteration if $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 \leq 0$

In figure 6.5, after the second iteration all the synchronization points lie in the third quadrant and all three streams are synchronized, since $(P_{3,1})_2$ lies in the third quadrant. However, in the situation

depicted in figure 6.6, the streams are not synchronized after the second iteration, since $(P_{3,1})_2$ lies in the fourth quadrant. Another static delay (β_3) equivalent to $(\gamma_{3,1})_2$ or $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0$ should be added to stream 3 to synchronize the streams. The value of the synchronization coefficients $\gamma_{1,2}$, $\gamma_{2,3}$ and $\gamma_{3,1}$ after the third iteration are given by,

$$(\gamma_{1,2})_3 = (\rho_2 + \beta_2) - (\rho_1 + \beta_1) + J^{\max}_2 - S_{1,2} \quad (6.27)$$

$$(\gamma_{2,3})_3 = (\rho_3 + \beta_3) - (\rho_2 + \beta_2) + J^{\max}_3 - S_{2,3} \quad (6.28)$$

$$(\gamma_{3,1})_3 = (\rho_1 + \beta_1) - (\rho_3 + \beta_3) + J^{\max}_1 - S_{3,1} \quad (6.29)$$

By substituting values

$$(\gamma_{1,2})_3 = (\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 \quad (6.30)$$

$$(\gamma_{2,3})_3 = 0 \quad (6.31)$$

$$(\gamma_{3,1})_3 = 0 \quad (6.32)$$

As $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 > 0$ point $(P_{1,2})_3$ lies in the fourth quadrant.

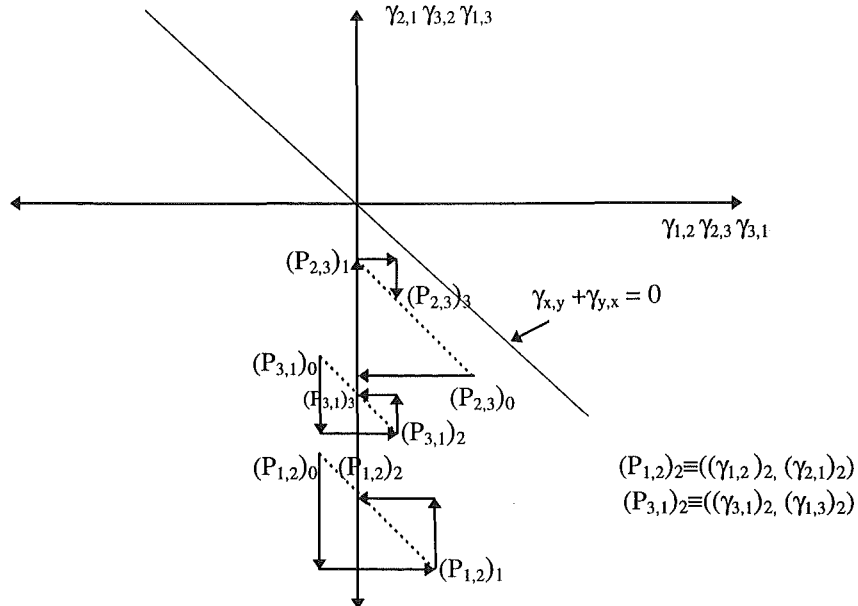


Figure 6.6 : Cartesian diagram showing the second iteration if $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 > 0$

Now another static delay needs to be added and the values of the synchronization coefficients would be

$$(\gamma_{1,2})_4 = 0 \quad (6.33)$$

$$(\gamma_{2,3})_4 = (\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 \quad (6.34)$$

$$(\gamma_{3,1})_4 = 0 \quad (6.35)$$

It can be seen that the value $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0$ will alternate between the synchronization coefficients $\gamma_{1,2}$, $\gamma_{2,3}$ and $\gamma_{3,1}$, if $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,1})_0 > 0$, and the streams cannot be synchronized. Similarly it can be shown that, if $(\gamma_{1,3})_0 + (\gamma_{3,2})_0 + (\gamma_{2,1})_0 > 0$, the streams cannot be synchronized. Thus, to synchronize the three media streams the following constraints should be satisfied :

$$\gamma_{1,2} + \gamma_{2,1} \leq 0 \text{ and}$$

$$\gamma_{1,3} + \gamma_{3,1} \leq 0 \text{ and}$$

$$\gamma_{2,3} + \gamma_{3,2} \leq 0 \text{ and}$$

$$\gamma_{1,2} + \gamma_{2,3} + \gamma_{3,1} \leq 0 \text{ and}$$

$$\gamma_{1,3} + \gamma_{3,2} + \gamma_{2,1} \leq 0.$$

where all synchronization coefficients refer to the zeroth iteration.

6.3 Derivation of Static Delay Value for Three Streams

The static delay that should be added to any stream to synchronize it with the other two streams depends on the synchronization coefficients. In this section, first, all the cases of synchronization coefficient values are studied and then an equation is derived to calculate the static delay to be added to stream 2 to synchronize it with the other two streams, assuming that all the synchronization constraints are satisfied. As stream 1 and 3 are analogous with stream 2, the above equation could be directly applied to calculate the static delays of the other two streams.

The static delay that should be added to stream 2 (β_2) is directly dependent on synchronization coefficients $(\gamma_{2,1})_0$ and $(\gamma_{2,3})_0$. The value β_2 is indirectly dependent on synchronization coefficients $(\gamma_{3,1})_0$ and $(\gamma_{1,3})_0$, as a static delay that is added to stream 1 or stream 3 may disturb the synchronization of stream 2. However, the value β_2 is independent of the synchronization coefficients $(\gamma_{1,2})_0$ and $(\gamma_{3,2})_0$ as these coefficients represent static delays to be added to stream 1

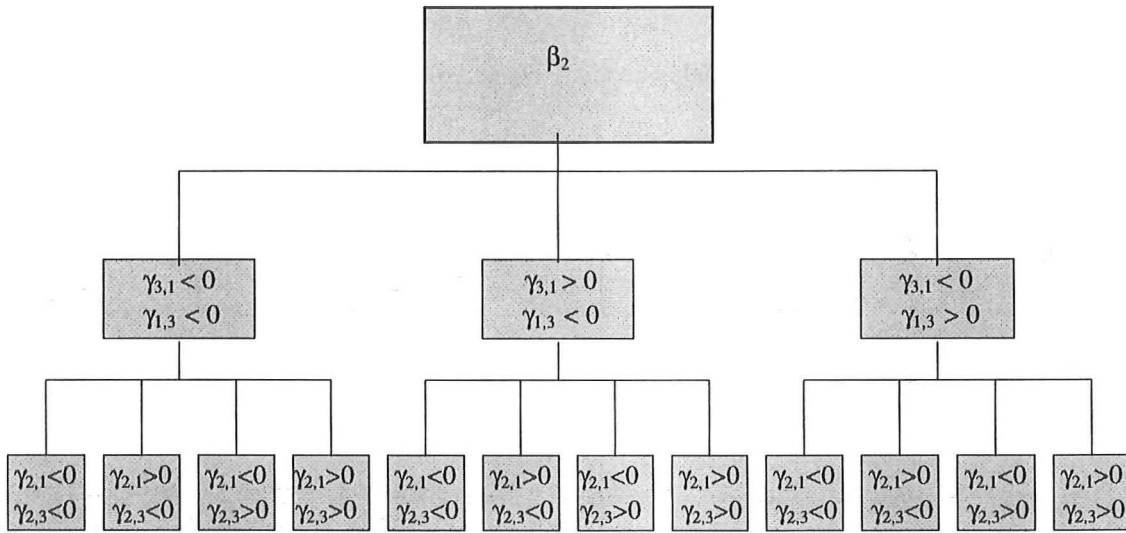
and stream 3 respectively to synchronize it with stream 2. The different cases that need to be considered are depicted by the chart in figure 6.7. If both $(\gamma_{3,1})_0$ and $(\gamma_{1,3})_0$ are positive then streams 1 and 3 are not able to be synchronized.

The following cases need to be investigated when considering $(\gamma_{3,1})_0$ and $(\gamma_{1,3})_0$.

1. Both $(\gamma_{3,1})_0$ and $(\gamma_{1,3})_0$ are negative
2. $(\gamma_{3,1})_0$ is positive while $(\gamma_{1,3})_0$ is negative
3. $(\gamma_{1,3})_0$ is positive while $(\gamma_{3,1})_0$ is negative

The following cases need to be investigated when considering $(\gamma_{2,1})_0$ and $(\gamma_{2,3})_0$.

1. Both $(\gamma_{2,1})_0$ and $(\gamma_{2,3})_0$ are negative
2. $(\gamma_{2,1})_0$ is positive while $(\gamma_{2,3})_0$ is negative.
3. $(\gamma_{2,3})_0$ is positive while $(\gamma_{2,1})_0$ is negative.
4. Both $(\gamma_{2,1})_0$ and $(\gamma_{2,3})_0$ are positive



case 1 case 2a case 2b case 3 case 4a case 5a case 5b case 6a case 4b case 5c case 5d case 6b

Figure 6.7 : The different cases when deriving β_2 (All synchronization coefficient values belong to the zeroth iteration)

Altogether there are twelve cases that need to be considered. These can be reduced to six representative cases as shown in figure 6.8. These are marked as case 1 to case 6, in figure 6.6. The result for other cases could be inferred from these six representative cases. In the next few subsections the above mentioned cases are discussed in detail, and finally an equation is derived for the static delay that should be added to stream 2.

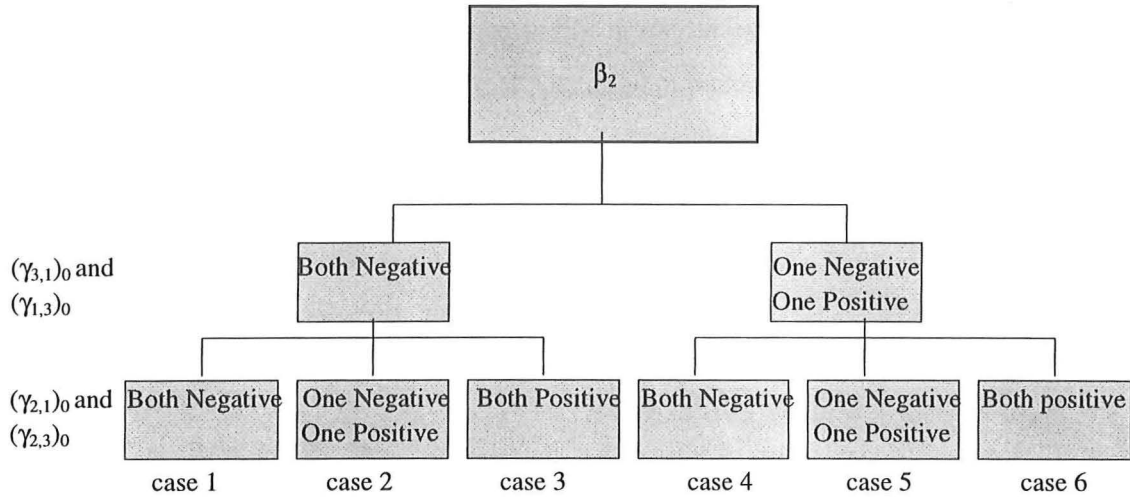


Figure 6.8 : A summary of the different representative cases when deriving β_2

6.3.1 Case 1 : $(\gamma_{2,1})_0 < 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 < 0$ and $(\gamma_{1,3})_0 < 0$

One of the possible situations of this case, is where all three synchronization points lie in the third quadrant, as shown in figure 6.9. Here, all three streams are synchronized with each other and the static delay to be added to stream 2, is zero. In the situation depicted in figure 6.10, $(P_{1,2})_0$ lies in the fourth quadrant, ie. $(\gamma_{1,2})_0 > 0$, and a static delay equivalent to $(\gamma_{1,2})_0$ should be added to stream 1, but the static delay to be added to stream 2, is zero. Similarly, if $(P_{2,3})_0$ lies in the second quadrant, ie. $(\gamma_{3,2})_0 > 0$, no static delay needs to be added to stream 2.

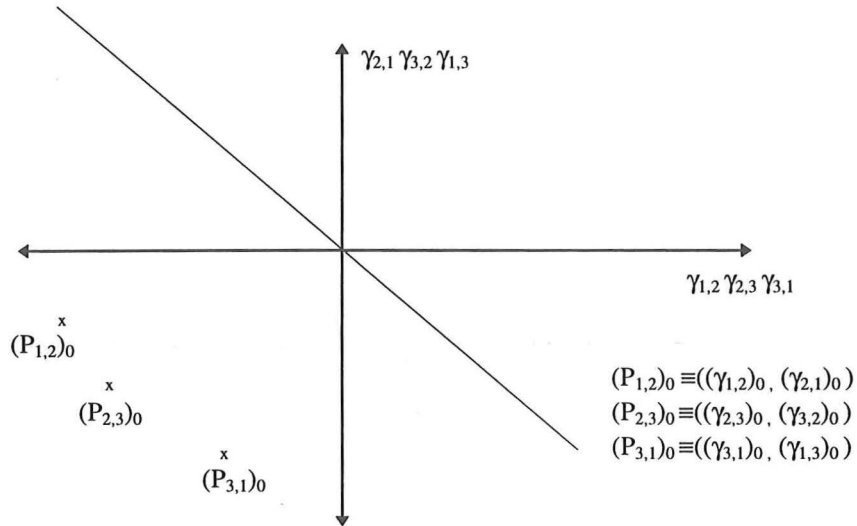


Figure 6.9 : All three streams are synchronized already

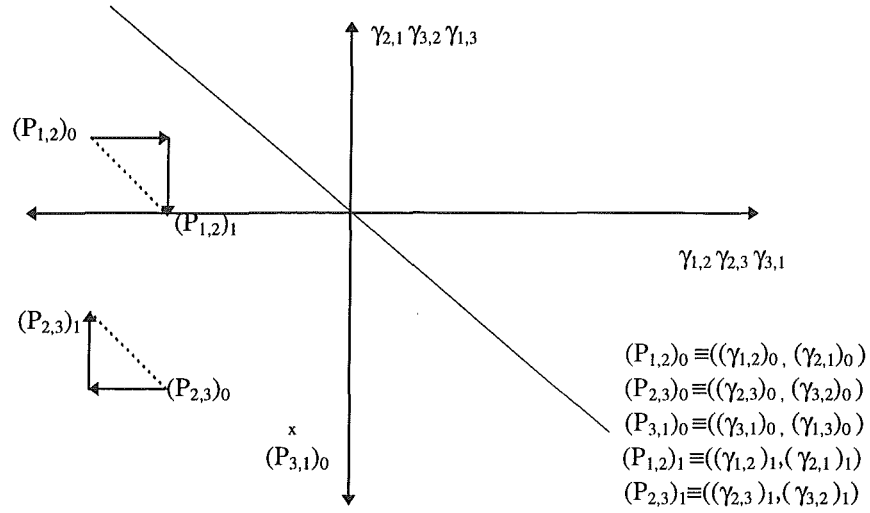


Figure 6.11 : Synchronization coefficients of the three streams $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 < 0$ and $(\gamma_{1,3})_0 < 0$
 $(\gamma_{3,2})_0 + (\gamma_{2,1})_0 < 0$

$(\gamma_{3,1})_0$ and $(\gamma_{1,3})_0$ are unaffected as the static delay is added to stream 2. In figure 6.11, after the first iteration all the points lie in the third quadrant and hence the streams are synchronized. In the situation shown in figure 6.12, after the addition of the static delay to stream 2, the synchronization between stream 2 and stream 3 is affected and to rectify this another static delay needs to be added to stream 3. However, the static delay of stream 2 remains equivalent to $(\gamma_{2,1})_0$.

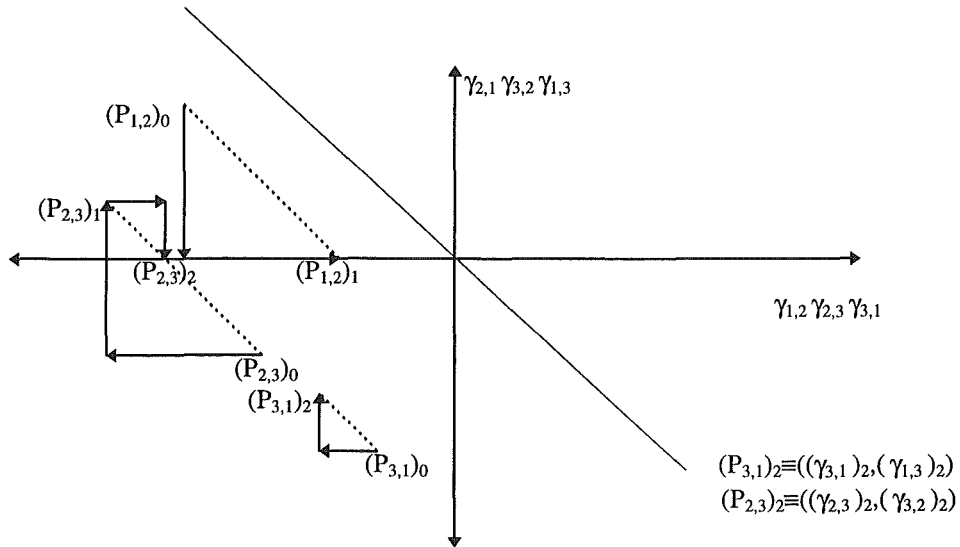


Figure 6.12 : Synchronization coefficients of the three streams $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 < 0$ and $(\gamma_{1,3})_0 < 0$
 $(\gamma_{3,2})_0 + (\gamma_{2,1})_0 > 0$

6.3.3 Case 3 : $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 < 0$ and $(\gamma_{1,3})_0 < 0$

Figure 6.13 depicts a situation where both $(\gamma_{2,1})_0 > 0$ and $(\gamma_{2,3})_0 > 0$. A static delay equivalent to $(\gamma_{2,1})_0$ have to be added to stream 2 to synchronize it with stream 1 and a static delay equivalent to $(\gamma_{2,3})_0$ have to be added to stream 2 to synchronize it with stream 3. If a static delay equivalent to $(\gamma_{2,1})_0$ is added to stream 2 the corresponding synchronization coefficient values after the first iteration is denoted by subscript x and if a static delay equivalent to $(\gamma_{2,3})_0$ is added to stream 2 the corresponding synchronization coefficient values after the first iteration is denoted by subscript y.

The synchronization point $(P_{2,3})_x$ lies in the fourth quadrant and the resultant streams are not synchronized. However, it is seen that all the synchronization points denoted by subscript y, lie in the third quadrant and the resultant streams are synchronized. Therefore the static delay that should be added to stream 2 is given by $(\gamma_{2,3})_0$ in this situation.

In figure 6.14 a reverse scenario is presented. Here, the static delay that should be added to stream 2 is equivalent to $(\gamma_{2,1})_0$. The two situations that are represented by figure 6.13 and figure 6.14 can be summarized as follows. When $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 < 0$ and $(\gamma_{1,3})_0 < 0$ the static delay that should be added to stream 2 is represented by $\max((\gamma_{2,1})_0, (\gamma_{2,3})_0)$.

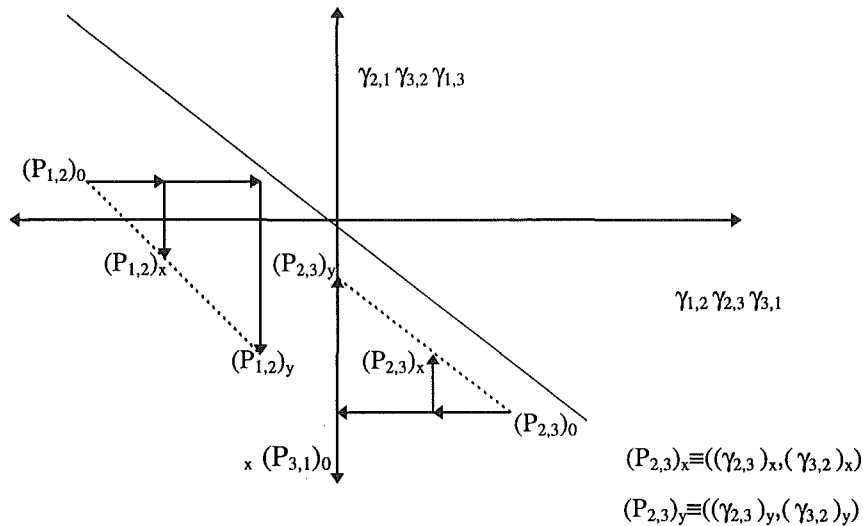


Figure 6.13 : Synchronization coefficients of the three streams $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 < 0$, $(\gamma_{1,3})_0 < 0$ and $(\gamma_{2,3})_0 > (\gamma_{2,1})_0$

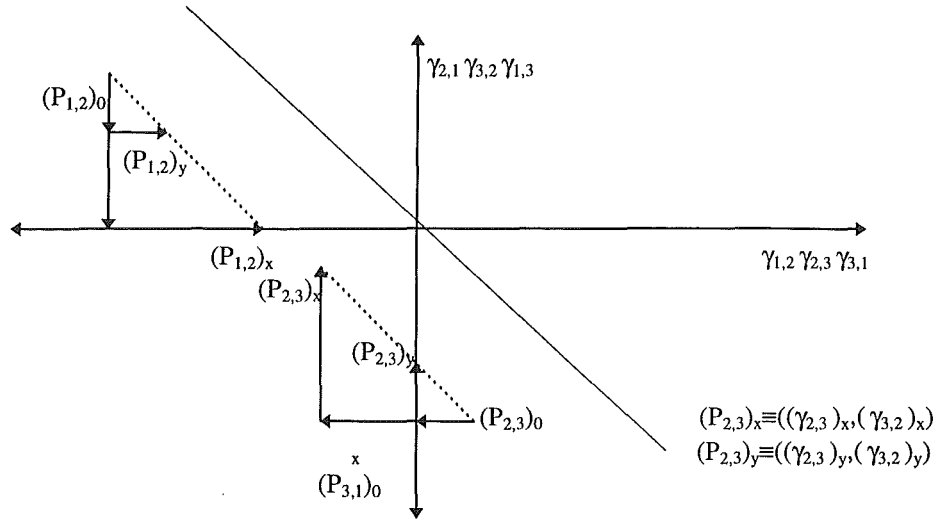


Figure 6.14 : Synchronization coefficients of the three streams $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 < 0$, $(\gamma_{1,3})_0 < 0$ and $(\gamma_{2,3})_0 < (\gamma_{2,1})_0$

6.3.4 Case 4 : $(\gamma_{2,1})_0 < 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 > 0$ and $(\gamma_{1,3})_0 < 0$

In the situation shown in figure 6.15, only $(\gamma_{3,1})_0$ is positive and therefore a static delay equivalent to $(\gamma_{3,1})_0$ should be added to stream 3 to synchronize it with stream 1. No other static delay needs to be added as all the points lie in the third quadrant after the first iteration. The point $(P_{2,3})_1$ would lie in the third quadrant only if the factor $(\gamma_{2,3})_0 + (\gamma_{3,1})_0 < 0$. Therefore, if $(\gamma_{2,3})_0 + (\gamma_{3,1})_0 < 0$ then no static delay needs to be added to stream 2.

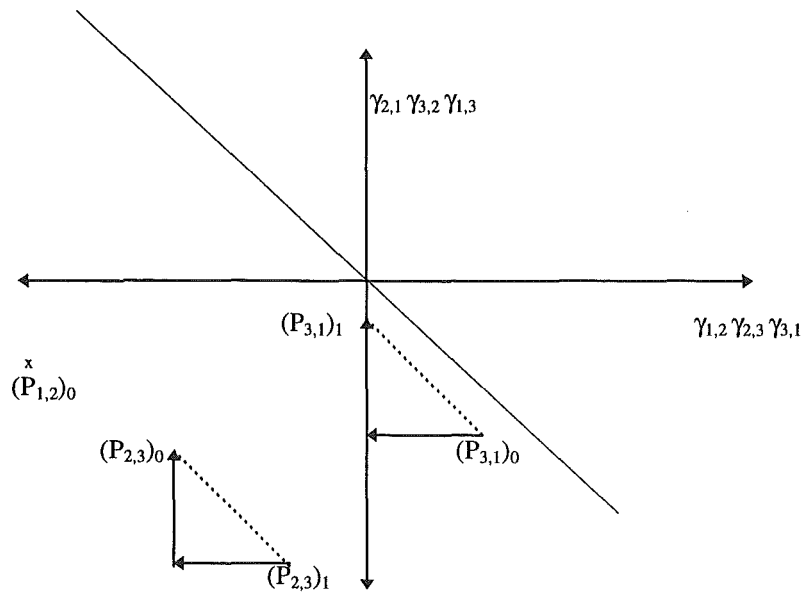


Figure 6.15 : Synchronization coefficients for three streams $(\gamma_{2,1})_0 < 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 > 0$, $(\gamma_{1,3})_0 < 0$ and $(\gamma_{2,3})_0 + (\gamma_{3,1})_0 < 0$

6.3.5 Case 5 : $(\gamma_{2,1})_0 > 0, (\gamma_{2,3})_0 < 0, (\gamma_{3,1})_0 > 0$ and $(\gamma_{1,3})_0 < 0$

In figure 6.17, the effective β_2 value is equivalent to $(\gamma_{2,1})_0$, for synchronization while in figure 6.18, the effective β_2 value for synchronization is equivalent to $(\gamma_{2,3})_0 + (\gamma_{3,1})_0$. Summarizing the situations depicted in figures 6.17 and 6.18 it can be said that when $(\gamma_{3,1})_0 > 0$ and $(\gamma_{2,1})_0 > 0$, $\beta_2 = \max((\gamma_{2,1})_0, (\gamma_{2,3})_0 + (\gamma_{3,1})_0)$.

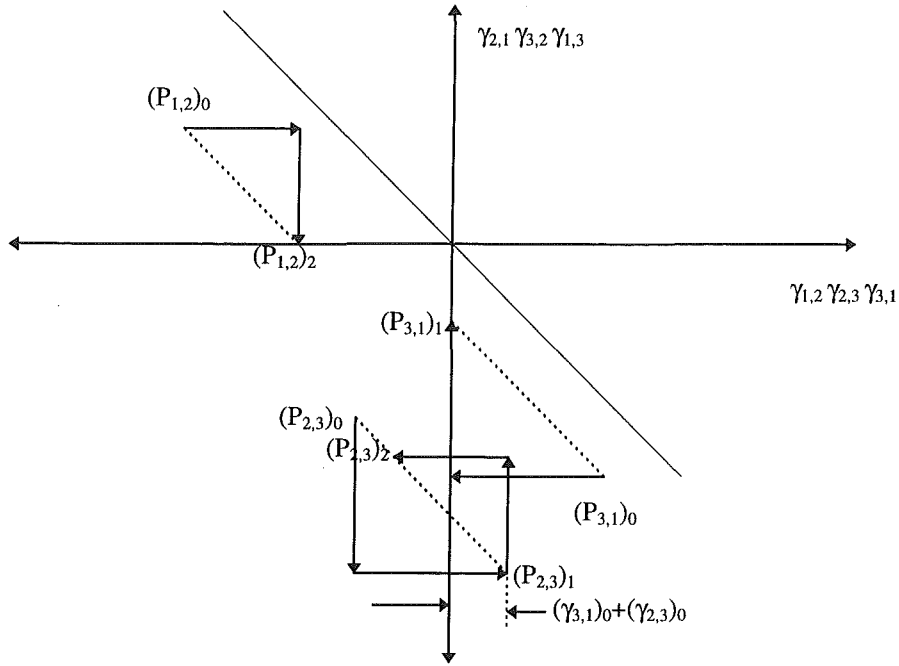


Figure 6.17 : Synchronization coefficient values for three streams for $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 > 0$, $(\gamma_{1,3})_0 > 0$ and $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 < (\gamma_{2,1})_0$.

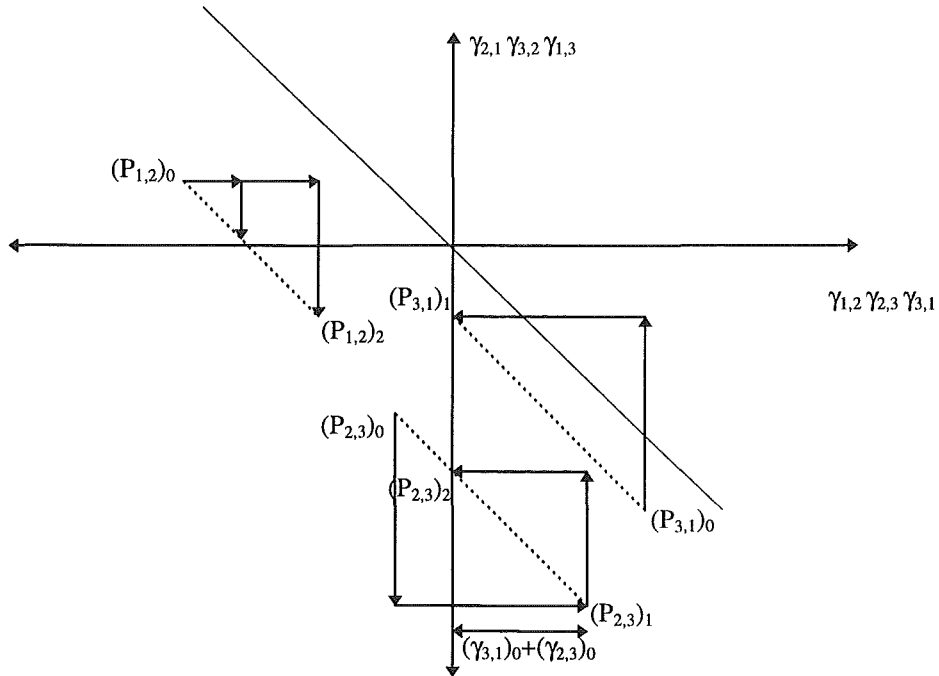


Figure 6.18 : Synchronization coefficient values for three streams for $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 < 0$, $(\gamma_{3,1})_0 > 0$, $(\gamma_{1,3})_0 > 0$ and $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 > (\gamma_{2,1})_0$.

6.3.6 Case 6 : $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 > 0$ and $(\gamma_{1,3})_0 < 0$

As shown in figure 6.19, A static delay equivalent to $(\gamma_{3,1})_0$ is added to stream 3 and the static delay, that should be added to stream 2, is equal to $(\gamma_{3,1})_0 + (\gamma_{2,3})_0$ as $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 > (\gamma_{2,1})_0$. However, in the situation shown in figure 6.20, the static delay that should be added to stream 2 is equal to $(\gamma_{2,1})_0$ as $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 < (\gamma_{2,1})_0$. In summary, if all three coefficients $\gamma_{3,1}$, $\gamma_{2,3}$ and $\gamma_{2,1}$ are positive then $\beta_2 = \max ((\gamma_{2,1})_0 ; (\gamma_{2,3})_0 + (\gamma_{3,1})_0)$.

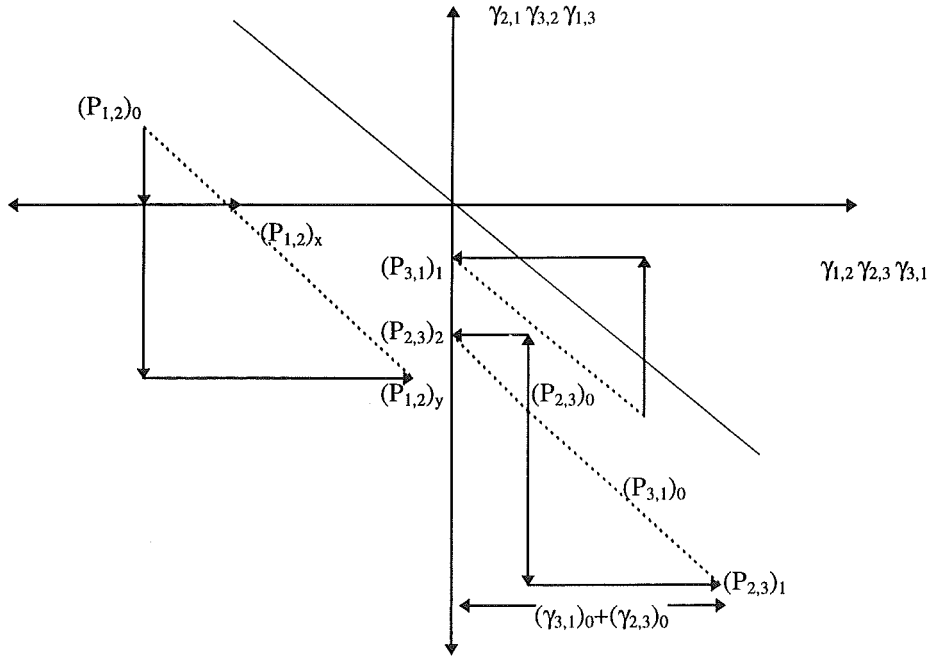


Figure 6.19 : Synchronization coefficient values for three streams for $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 > 0$, $(\gamma_{1,3})_0 < 0$ and $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 > (\gamma_{2,1})_0$.

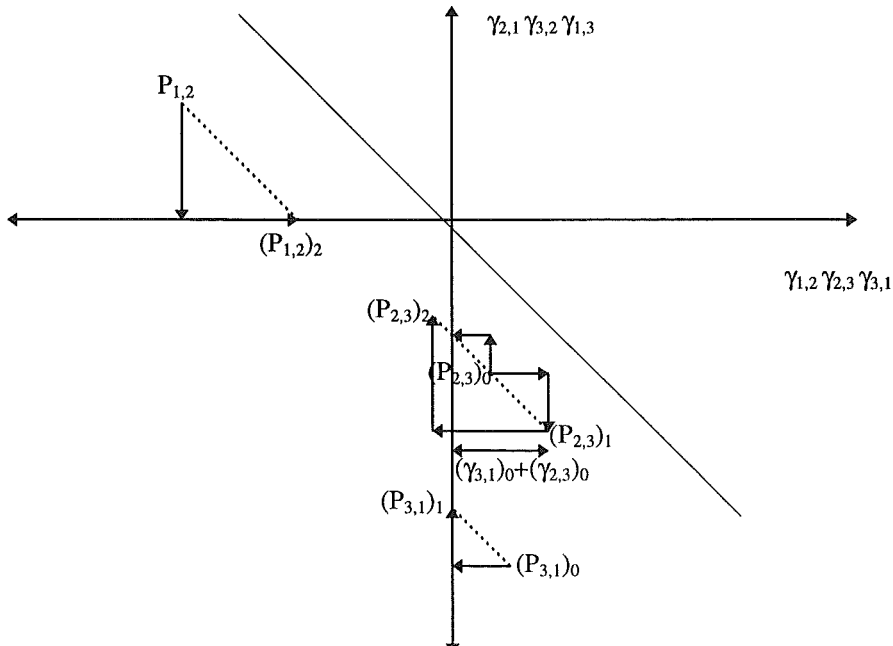


Figure 6.20 : Static delay coefficient values for three streams for $(\gamma_{2,1})_0 > 0$, $(\gamma_{2,3})_0 > 0$, $(\gamma_{3,1})_0 > 0$, $(\gamma_{1,3})_0 < 0$ and $(\gamma_{3,1})_0 + (\gamma_{2,3})_0 < (\gamma_{2,1})_0$.

6.3.7 Summary of All Possible Situations

Figure 6.21 summarizes static delay of stream 2 (β_2) for the six representative cases. Table 6.1 summarizes the β_2 values for all 12 possible cases. The β_2 values for the other six cases are inferred.

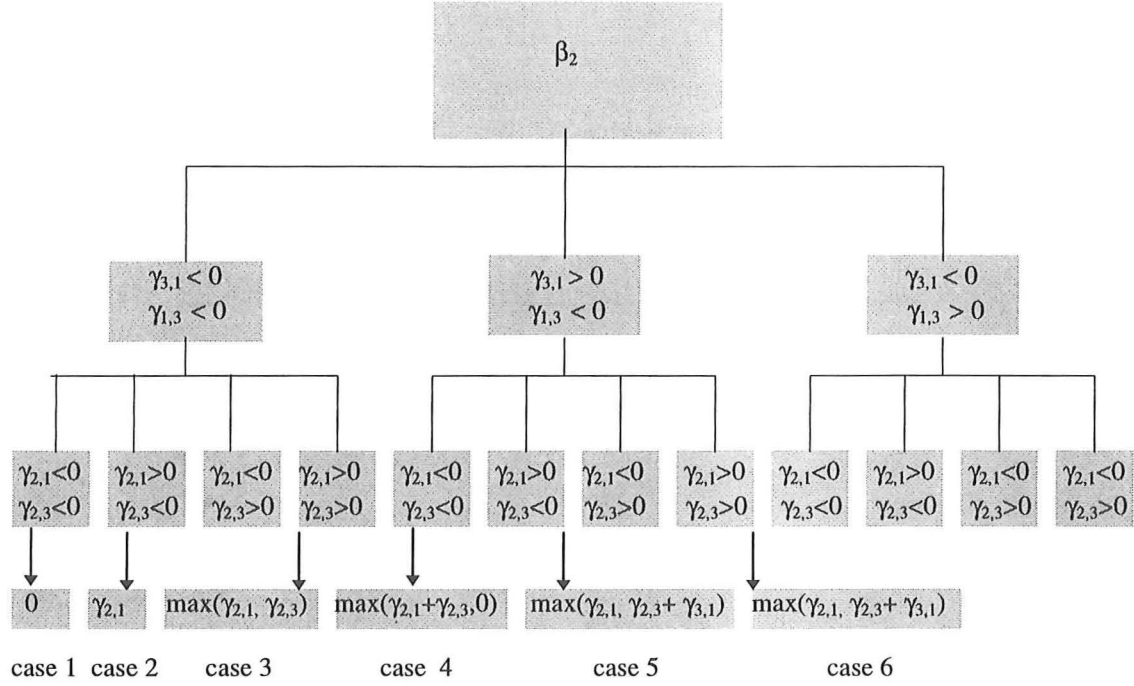


Figure 6.21 : β_2 values for the six representative cases (all synchronization coefficients belong to the zeroth iteration or initial condition)

Case	$\gamma_{3,1}$	$\gamma_{1,3}$	$\gamma_{2,1}$	$\gamma_{2,3}$	β_2	$\beta_2(\text{summarized})$	Remarks
1	$\gamma_{3,1} < 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} < 0$	0	$\max(\gamma_{2,1}, \gamma_{2,3}, 0)$	case 1
2a	$\gamma_{3,1} < 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} < 0$	$\gamma_{2,1}$		case 2
2b	$\gamma_{3,1} < 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} > 0$	$\gamma_{2,3}$		inferred
3	$\gamma_{3,1} < 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} > 0$	$\max(\gamma_{2,1}, \gamma_{2,3})$	$\max(\gamma_{2,1}, \gamma_{2,3} + \gamma_{3,1}, 0)$	case 3
4a	$\gamma_{3,1} > 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} < 0$	$\max(\gamma_{2,3} + \gamma_{3,1}, 0)$		case 4
5a	$\gamma_{3,1} > 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} < 0$	$\max(\gamma_{2,1}, \gamma_{2,3} + \gamma_{3,1})$		case 5
5b	$\gamma_{3,1} > 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} > 0$	$\gamma_{2,3} + \gamma_{3,1}$	$\max(\gamma_{2,1}, \gamma_{2,3} + \gamma_{3,1}, 0)$	inferred
6a	$\gamma_{3,1} > 0$	$\gamma_{1,3} < 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} > 0$	$\max(\gamma_{2,1}, \gamma_{2,3} + \gamma_{3,1})$		case 6
4b	$\gamma_{3,1} < 0$	$\gamma_{1,3} > 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} < 0$	$\max(\gamma_{2,1} + \gamma_{1,3}, 0)$	$\max(\gamma_{2,1} + \gamma_{1,3}, \gamma_{2,3}, 0)$	inferred
5c	$\gamma_{3,1} < 0$	$\gamma_{1,3} > 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} < 0$	$\gamma_{2,1} + \gamma_{1,3}$		inferred
5d	$\gamma_{3,1} < 0$	$\gamma_{1,3} > 0$	$\gamma_{2,1} < 0$	$\gamma_{2,3} > 0$	$\max(\gamma_{2,1} + \gamma_{1,3}, \gamma_{2,3})$		inferred
6b	$\gamma_{3,1} < 0$	$\gamma_{1,3} > 0$	$\gamma_{2,1} > 0$	$\gamma_{2,3} > 0$	$\max(\gamma_{2,1} + \gamma_{1,3}, \gamma_{2,3})$		inferred

Table 6.1 : Summary of static delay values of stream 2 (all synchronization coefficients belong to the zeroth iteration).

Note that the β_2 value for case 5b is inferred from case 5a. In case 5a, the β_2 value is equivalent to $\max(\gamma_{2,1}, \gamma_{2,3} + \gamma_{3,1})$. In case 5b, $\gamma_{2,1}$ is negative and both $\gamma_{2,3}$ and $\gamma_{3,1}$ are positive, and therefore the sum $\gamma_{2,3} + \gamma_{3,1}$ is always positive and greater than $\gamma_{2,1}$. Hence the value of β_2 for case 5b is equivalent to $\gamma_{2,3} + \gamma_{3,1}$. A similar argument holds for case 5c.

All the possible cases can be summarized as

$$\beta_2 = \max\{(\gamma_{2,1} + a_{1,3} * \gamma_{1,3}), (\gamma_{2,3} + a_{3,1} * \gamma_{3,1}), 0\}$$

$$\text{where } a_{3,1} = \begin{cases} 0 & \text{when } \gamma_{3,1} \leq 0 \\ 1 & \text{when } \gamma_{3,1} > 0 \end{cases}$$

$$\text{and } a_{1,3} = \begin{cases} 0 & \text{when } \gamma_{1,3} \leq 0 \\ 1 & \text{when } \gamma_{1,3} > 0 \end{cases}$$

In other words, the value of β_2 can be written as

$$\beta_2 = \max\{a_{2,1,3} * (\gamma_{2,1} + a_{1,3} * \gamma_{1,3}), a_{2,3,1} * (\gamma_{2,3} + a_{3,1} * \gamma_{3,1})\} \quad (6.44)$$

$$\text{where } a_{2,1,3} = \begin{cases} 0 & \text{when } \gamma_{2,1} + a_{1,3} * \gamma_{1,3} \leq 0 \\ 1 & \text{when } \gamma_{2,1} + a_{1,3} * \gamma_{1,3} > 0 \end{cases}$$

$$\text{and } a_{2,3,1} = \begin{cases} 0 & \text{when } \gamma_{2,3} + a_{3,1} * \gamma_{3,1} \leq 0 \\ 1 & \text{when } \gamma_{2,3} + a_{3,1} * \gamma_{3,1} > 0 \end{cases}$$

The above is a generalized equation for static delay of stream 2. Similarly generalized equations for static delays of stream 1 and 3 can be written as

$$\beta_1 = \max\{a_{1,2,3} * (\gamma_{1,2} + a_{2,3} * \gamma_{2,3}), a_{1,3,2} * (\gamma_{1,3} + a_{3,2} * \gamma_{3,2})\} \quad (6.45)$$

$$\beta_3 = \max\{a_{3,1,2} * (\gamma_{3,1} + a_{1,2} * \gamma_{1,2}), a_{3,2,1} * (\gamma_{3,2} + a_{2,1} * \gamma_{2,1})\} \quad (6.46)$$

where $a_{2,3}$, $a_{3,2}$, $a_{1,2}$ and $a_{2,1}$ are defined in a way to $a_{3,1}$ and $a_{1,2,3}$, $a_{1,3,2}$, $a_{3,1,2}$ and $a_{3,2,1}$ are defined in a similar way to $a_{2,1,3}$. In the next chapter, these equations are extended to calculate the static delay that should be added to any media stream, when synchronizing n media streams. Synchronization of layered coding systems and a few real valued examples are also discussed in the next chapter.

CHAPTER 7

STATIC DELAY COMPENSATION; A MECHANISM FOR SYNCHRONIZING MULTIPLE MEDIA SYSTEMS

In synchronizing multiple media streams, there are a number of constraints that need to be satisfied. In this chapter, an equation is derived to calculate the static delay that should be added to any media stream, when synchronizing n media streams. Synchronization of layered coding systems is also discussed. In order to synchronize the layered streams static delays are added to the streams with lower end-to-end delays. A few real valued examples of synchronization of multiple media streams are discussed at the end of the chapter.

7.1 Constraints to be Satisfied when Synchronizing n Media Streams

In synchronizing n media streams there are a multitude of constraints that need to be satisfied.

Considering different combinations of pairs to be synchronized,

$$\gamma_{1,2} + \gamma_{2,1} < 0, \gamma_{1,3} + \gamma_{3,1} < 0, \dots, \gamma_{1,n} + \gamma_{n,1} < 0, \dots, \gamma_{n,n-1} + \gamma_{n-1,n} < 0.$$

The number of constraints = number of combinations of 2 elements out of $n = nC_2$.

Any combination of three streams should also be able to be synchronized. Considering stream i, j and k $\gamma_{i,j} + \gamma_{j,k} + \gamma_{k,i} < 0$ and $\gamma_{j,i} + \gamma_{i,k} + \gamma_{k,j} < 0$ need to be satisfied (section 6.2).

There are nC_3 possible combinations of 3 streams and in each combination, two constraints need to be satisfied. The number of constraints = $2 * nC_3$.

Any combination of four streams should also be able to be synchronized. Considering stream i, j, k , and l , the constraints that need to be satisfied are,

$$\gamma_{i,j} + \gamma_{j,k} + \gamma_{k,l} + \gamma_{l,i} < 0, \quad \gamma_{i,j} + \gamma_{j,l} + \gamma_{l,k} + \gamma_{k,i} < 0,$$

$$\gamma_{i,k} + \gamma_{k,j} + \gamma_{j,l} + \gamma_{l,i} < 0, \quad \gamma_{i,k} + \gamma_{k,l} + \gamma_{l,j} + \gamma_{j,i} < 0,$$

$$\gamma_{i,l} + \gamma_{l,j} + \gamma_{j,k} + \gamma_{k,i} < 0, \quad \gamma_{i,l} + \gamma_{l,k} + \gamma_{k,j} + \gamma_{j,i} < 0.$$

There are nC_4 possible combinations of 4 streams and in each combination, six constraints need to be satisfied. The number of constraints = $6 * nC_4$.

⋮

Any combination of r streams should also be able to be synchronized. There are nC_r possible combinations of r streams and in each combination, $(r-1)!$ constraints need to be satisfied.

The number of constraints = $(r-1)! * nC_r$.

Therefore the total number of constraints that need to be satisfied = $\sum_{r=1}^n (r-1)! * nC_r$. If all the above constraints are satisfied then, **the streams are able to be synchronized.**

7.2 Static Delay Value for n Media Streams

In this section, the equation for static delay, when synchronizing three media streams is extended to calculate the static delay value of any media stream when synchronizing n media streams. When synchronizing three media streams, the static delays to be added to three media streams are given by (section 6.3.7),

$$\beta_1 = \max \{ a_{1,2,3} * (\gamma_{1,2} + a_{2,3} * \gamma_{2,3}), a_{1,3,2} * (\gamma_{1,3} + a_{3,2} * \gamma_{3,2}) \} \quad (7.1)$$

$$\beta_2 = \max \{ a_{2,1,3} * (\gamma_{2,1} + a_{1,3} * \gamma_{1,3}), a_{2,3,1} * (\gamma_{2,3} + a_{3,1} * \gamma_{3,1}) \} \quad (7.2)$$

$$\beta_3 = \max \{ a_{3,1,2} * (\gamma_{3,1} + a_{1,2} * \gamma_{1,2}), a_{3,2,1} * (\gamma_{3,2} + a_{2,1} * \gamma_{2,1}) \} \quad (7.3)$$

The static delay to be added to stream 1 (β_1) is directly dependent on synchronization coefficients $\gamma_{1,2}$ and $\gamma_{1,3}$, and indirectly dependent on the coefficients $\gamma_{2,3}$, $\gamma_{3,2}$. The indirectly dependent coefficients are active only if they are positive. The value β_1 is independent of synchronization coefficients $\gamma_{2,1}$ and $\gamma_{3,1}$.

When synchronizing four media streams, the value β_1 is directly dependent on synchronization coefficients $\gamma_{1,2}$, $\gamma_{1,3}$, and $\gamma_{1,4}$, and indirectly dependent on the coefficients $\gamma_{2,3}$, $\gamma_{3,2}$, $\gamma_{2,4}$, $\gamma_{4,2}$, $\gamma_{3,4}$ and $\gamma_{4,3}$. The indirectly dependent coefficients are active only if they are positive. The value β_1 is independent of synchronization coefficients $\gamma_{2,1}$, $\gamma_{3,1}$ and $\gamma_{4,1}$.

Consider the synchronization coefficients $\gamma_{1,2}$, $\gamma_{2,3}$, $\gamma_{3,4}$ and $\gamma_{4,1}$ as depicted in figure 7.1, which represents the worst case, where three iterations are necessary to achieve synchronization. As synchronization coefficient $(\gamma_{3,4})_0$ is positive, a static delay should be added to stream 3 to synchronize it with stream 4. Also as the sum $(\gamma_{2,3})_0 + (\gamma_{3,4})_0$ is positive the synchronization between stream 2 and 3 is disturbed, and a static delay of $(\gamma_{2,3})_0 + (\gamma_{3,4})_0$ should be added to stream 2 to re-synchronize it with stream 3. As a consequence of the above, synchronization between stream 1 and 2 is disturbed since the sum $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,4})_0$ is positive, and another static delay equivalent to $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,4})_0$ should be added to stream 1, to synchronize it with stream 2. The above procedure cannot affect the synchronization between stream 1 and 4 because the sum $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,4})_0 + (\gamma_{4,1})_0$ should be less than zero, if the media streams are able to be synchronized.

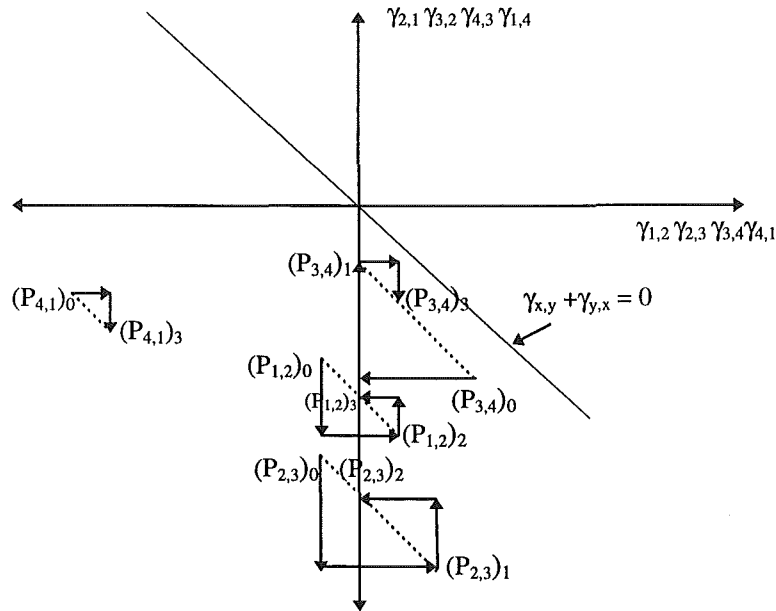


Figure 7.1 : Synchronizing four media streams when $(\gamma_{3,4})_0 > 0$, $(\gamma_{2,3})_0 + (\gamma_{3,4})_0 > 0$ and $(\gamma_{1,2})_0 + (\gamma_{2,3})_0 + (\gamma_{3,4})_0 > 0$

Thus, when considering synchronization coefficients $\gamma_{1,2}$, $\gamma_{2,3}$ and $\gamma_{3,4}$ only, static delay of stream 1 is equivalent to $a_{1,2,3,4} * (\gamma_{1,2} + a_{2,3,4} * (\gamma_{2,3} + a_{3,4} * \gamma_{3,4}))$, where

$$a_{1,2,3,4} = \begin{cases} 0 & \text{when } \gamma_{1,2} + a_{2,3,4} * (\gamma_{2,3} + a_{3,4} * \gamma_{3,4}) \leq 0 \\ 1 & \text{when } \gamma_{1,2} + a_{2,3,4} * (\gamma_{2,3} + a_{3,4} * \gamma_{3,4}) > 0 \end{cases}$$

Similarly all the possible permutations of the four streams starting from 1, need to be considered to find the overall static delay that should be added to stream 1. Thus, the static delay of stream 1 can be written as,

$$\beta_1 = \max \{ a_{1,2,3,4} * (\gamma_{1,2} + a_{2,3,4} * (\gamma_{2,3} + a_{3,4} * \gamma_{3,4})) , a_{1,2,4,3} * (\gamma_{1,2} + a_{2,4,3} * (\gamma_{2,4} + a_{4,3} * \gamma_{4,3})), \\ a_{1,3,2,4} * (\gamma_{1,3} + a_{3,2,4} * (\gamma_{3,2} + a_{2,4} * \gamma_{2,4})) , a_{1,3,4,2} * (\gamma_{1,3} + a_{3,4,2} * (\gamma_{3,4} + a_{4,2} * \gamma_{2,4})), \\ a_{1,4,2,3} * (\gamma_{1,4} + a_{4,2,3} * (\gamma_{4,2} + a_{2,3} * \gamma_{2,3})) , a_{1,4,3,2} * (\gamma_{1,4} + a_{4,3,2} * (\gamma_{4,3} + a_{3,2} * \gamma_{3,2})) \} , \quad (7.4)$$

where the coefficients $a_{2,4,3}$, $a_{3,2,4}$, $a_{3,4,2}$, $a_{4,2,3}$ and $a_{4,3,2}$ are defined in a similar way to $a_{2,3,4}$. Similarly, the static delays of the other three streams can be written as,

$$\beta_2 = \max \{ a_{2,1,3,4} * (\gamma_{2,1} + a_{1,3,4} * (\gamma_{1,3} + a_{3,4} * \gamma_{3,4})) , a_{2,1,4,3} * (\gamma_{2,1} + a_{1,4,3} * (\gamma_{1,4} + a_{4,3} * \gamma_{4,3})), \\ a_{2,3,1,4} * (\gamma_{2,3} + a_{3,1,4} * (\gamma_{3,1} + a_{1,4} * \gamma_{1,4})) , a_{2,3,4,1} * (\gamma_{2,3} + a_{3,4,1} * (\gamma_{3,4} + a_{4,1} * \gamma_{4,1})), \\ a_{2,4,1,3} * (\gamma_{2,4} + a_{4,1,3} * (\gamma_{4,1} + a_{1,3} * \gamma_{1,3})) , a_{2,4,3,1} * (\gamma_{2,4} + a_{4,3,1} * (\gamma_{4,3} + a_{3,1} * \gamma_{3,1})) \} \quad (7.5)$$

$$\beta_3 = \max \{ a_{3,1,2,4} * (\gamma_{3,1} + a_{1,2,4} * (\gamma_{1,2} + a_{2,4} * \gamma_{2,4})) , a_{3,1,4,2} * (\gamma_{3,1} + a_{1,4,2} * (\gamma_{1,4} + a_{4,2} * \gamma_{4,2})), \\ a_{3,2,1,4} * (\gamma_{3,2} + a_{2,1,4} * (\gamma_{2,1} + a_{1,4} * \gamma_{1,4})) , a_{3,2,4,1} * (\gamma_{3,2} + a_{2,4,1} * (\gamma_{2,4} + a_{4,1} * \gamma_{4,1})), \\ a_{3,4,1,2} * (\gamma_{3,4} + a_{4,1,2} * (\gamma_{4,1} + a_{1,2} * \gamma_{1,2})) , a_{3,4,2,1} * (\gamma_{3,4} + a_{4,2,1} * (\gamma_{4,2} + a_{2,1} * \gamma_{2,1})) \} \quad (7.6)$$

$$\beta_4 = \max \{ a_{4,1,2,3} * (\gamma_{4,1} + a_{1,2,3} * (\gamma_{1,2} + a_{2,3} * \gamma_{2,3})) , a_{4,1,3,2} * (\gamma_{4,1} + a_{1,3,2} * (\gamma_{1,3} + a_{3,2} * \gamma_{3,2})), \\ a_{4,2,1,3} * (\gamma_{4,2} + a_{2,1,3} * (\gamma_{2,1} + a_{1,3} * \gamma_{1,3})) , a_{4,2,3,1} * (\gamma_{4,2} + a_{2,3,1} * (\gamma_{2,3} + a_{3,1} * \gamma_{3,1})), \\ a_{4,3,1,2} * (\gamma_{4,3} + a_{3,1,2} * (\gamma_{3,1} + a_{1,2} * \gamma_{1,2})) , a_{4,3,2,1} * (\gamma_{4,3} + a_{3,2,1} * (\gamma_{3,2} + a_{2,1} * \gamma_{2,1})) \} \quad (7.7)$$

When synchronizing n media streams, the static delay to be added to stream 1 (β_1), all the permutations of n streams starting from 1 need to be considered. Therefore, the static delay of stream 1, can be written as

$$\beta_1 = \max \{ a_{1,2,3,\dots,n-1,n} (\gamma_{1,2} + a_{2,3,\dots,n} * (\gamma_{2,3} + a_{3,4,\dots,n} * (\gamma_{3,4} + \dots + a_{n-2,n-1,n} * (\gamma_{n-2,n-1} + a_{n-1,n} * \gamma_{n-1,n}))), \\ a_{1,2,3,\dots,n-1,n} (\gamma_{1,2} + a_{2,3,\dots,n} * (\gamma_{2,3} + a_{3,4,\dots,n} * (\gamma_{3,4} + \dots + a_{n-2,n-1,n} * (\gamma_{n-2,n} + a_{n,n-1} * \gamma_{n,n-1}))), \\ \vdots \\ a_{1,p,q,r,\dots,s,t,u} (\gamma_{1,p} + a_{p,q,\dots,t,u} * (\gamma_{p,q} + a_{q,r,\dots,u} * (\gamma_{q,r} + \dots + a_{s,t,u} * (\gamma_{s,t} + a_{t,u} * \gamma_{t,u}))), \\ \vdots \\ a_{1,n,n-1,\dots,3,2} (\gamma_{1,n} + a_{n,n-1,\dots,2} * (\gamma_{n,n-1} + a_{n-1,n-2,\dots,2} * (\gamma_{n-1,n-2} + \dots + a_{4,3,2} * (\gamma_{4,3} + a_{3,2} * \gamma_{3,2})))) \} \quad (7.8)$$

The above equation has $(n-1)!$ terms as it should encounter all possible permutations of $n-1$ elements. Similarly equations for other static delays can be written. In equation 7.8, all the coefficients are defined in a similar way to

$$a_{1,2,3,\dots,n-1,n} = \begin{cases} 0 & \text{when } \gamma_{1,2} + a_{2,3,\dots,n} * (\gamma_{2,3} + a_{3,4,\dots,n} * (\gamma_{3,4} + \dots + a_{n-2,n-1,n} * (\gamma_{n-2,n-1} + a_{n-1,n} * \gamma_{n-1,n}))) \leq 0 \\ 1 & \text{when } \gamma_{1,2} + a_{2,3,\dots,n} * (\gamma_{2,3} + a_{3,4,\dots,n} * (\gamma_{3,4} + \dots + a_{n-2,n-1,n} * (\gamma_{n-2,n-1} + a_{n-1,n} * \gamma_{n-1,n}))) > 0 \end{cases}$$

7.3 Example of Synchronizing Multiple Audio and Video Streams

This section considers a real valued example for synchronizing multiple media streams. In a multimedia conferencing application video and audio streams of different sources may have to be played out in synchrony at a destination. Consider the multimedia conferencing application shown in figure 7.2.

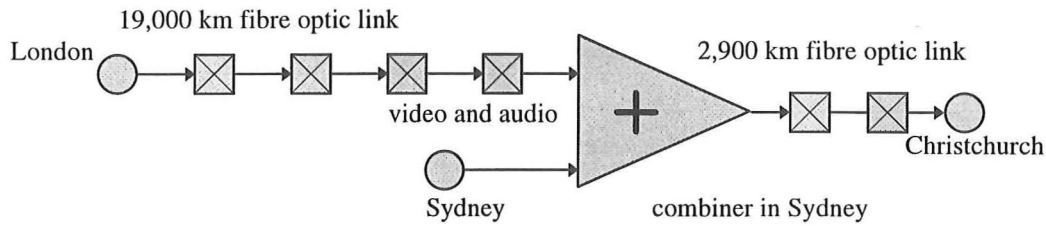


Figure 7.2 : Synchronizing multiple media streams : A pair each of video and audio streams originating from London and Sydney to be played out in synchrony at Christchurch.

The delays and jitters of the four streams are depicted in figure 7.3. At playout, video is allowed to lead audio by 90 ms and audio is allowed to lead video by 60 ms. Therefore, $S_{1,2} = 90$, $S_{2,1} = 60$. and $S_{3,4} = 90$, $S_{4,3} = 60$. Also it is important to have tight synchronization between the two audio streams. Therefore a value of 100 ms is chosen for both $S_{2,4}$ and $S_{4,2}$. There is a little significance in tightly synchronizing a video stream originating from a particular source with an audio stream of another source. However they should not be fully out of synchrony. Thus remaining skew tolerance values are taken to be equivalent to 400 ms. This is a problem of synchronizing four media streams. The calculated values of the synchronization coefficients are as follows.

$$\begin{array}{llll} \gamma_{1,2} = -151 & \gamma_{2,1} = 84.5 & \gamma_{1,3} = -461 & \gamma_{3,1} = -236 \\ \gamma_{1,4} = -563.75 & \gamma_{4,1} = -160.5 & \gamma_{2,3} = -385.5 & \gamma_{3,2} = -366 \\ \gamma_{2,4} = -188.25 & \gamma_{4,2} = 9.5 & \gamma_{3,4} = -158.75 & \gamma_{4,3} = 49.5 \end{array}$$

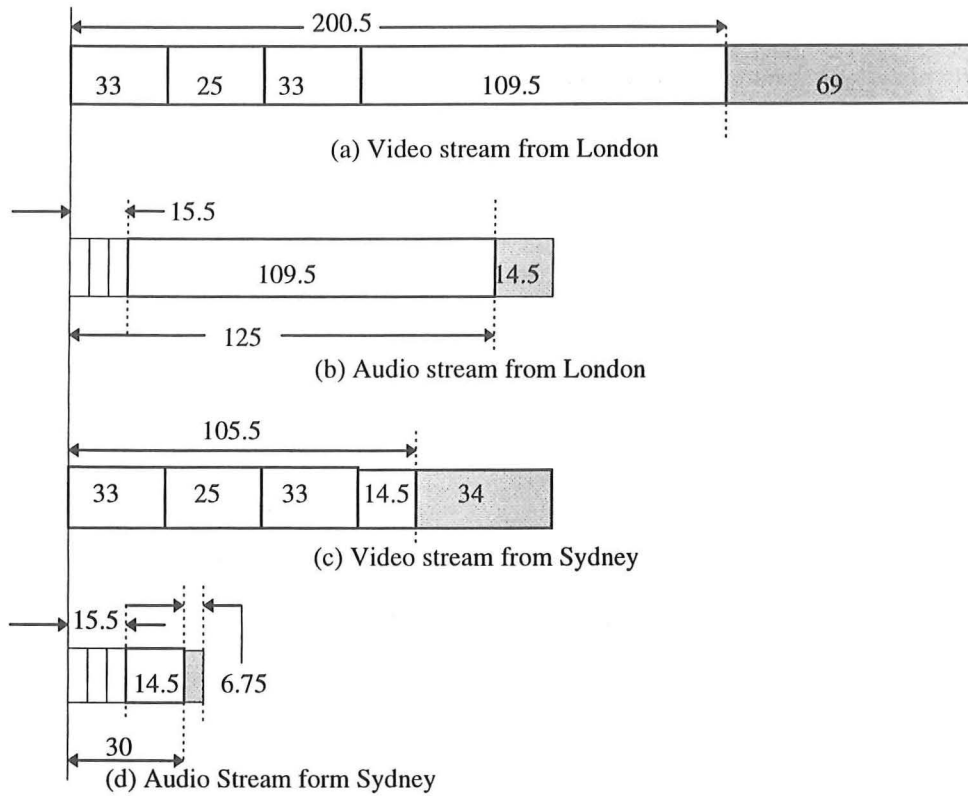


Figure 7.3 : Video and audio streams for situation depicted in figure 7.2

All the constraints have to be tested to see whether they are satisfied.

$$\gamma_{1,2} + \gamma_{2,1} = -66.5 \quad \gamma_{1,3} + \gamma_{3,1} = -697 \quad \gamma_{1,4} + \gamma_{4,1} = -724.25$$

$$\gamma_{2,3} + \gamma_{3,2} = -751.5 \quad \gamma_{2,4} + \gamma_{4,2} = -178.75 \quad \gamma_{3,4} + \gamma_{4,3} = -109.25$$

$$\gamma_{1,2} + \gamma_{2,3} + \gamma_{3,1} = -772.5 \quad \gamma_{1,2} + \gamma_{2,4} + \gamma_{4,1} = -499.75 \quad \gamma_{1,3} + \gamma_{3,2} + \gamma_{2,1} = -742.5$$

$$\gamma_{1,3} + \gamma_{3,4} + \gamma_{4,1} = -780.25 \quad \gamma_{1,4} + \gamma_{4,2} + \gamma_{2,1} = -469.75 \quad \gamma_{1,4} + \gamma_{4,3} + \gamma_{3,1} = -750.25$$

$$\gamma_{2,3} + \gamma_{3,4} + \gamma_{4,2} = -534.75 \quad \gamma_{2,4} + \gamma_{4,3} + \gamma_{3,2} = -504.75$$

$$\gamma_{1,2} + \gamma_{2,3} + \gamma_{3,4} + \gamma_{4,1} = -855.75 \quad \gamma_{1,3} + \gamma_{3,2} + \gamma_{2,4} + \gamma_{4,1} = -1175.75 \quad \gamma_{1,4} + \gamma_{4,3} + \gamma_{3,2} + \gamma_{2,1} = -795.75$$

$$\gamma_{1,2} + \gamma_{2,4} + \gamma_{4,3} + \gamma_{3,1} = -525.75 \quad \gamma_{1,3} + \gamma_{3,4} + \gamma_{4,2} + \gamma_{2,1} = -525.75 \quad \gamma_{1,4} + \gamma_{4,2} + \gamma_{2,3} + \gamma_{3,1} = -1175.75$$

Since all the above additions produce results that are less than zero, all the synchronization constraints are satisfied and the streams are able to be synchronized. Equations 7.4 to 7.7 should be used to calculate the static delays to be added to the four streams. By applying the above four equations, 7.4 to 7.7, $\beta_1 = 0$, $\beta_2 = 84.5$, $\beta_3 = 0$ and $\beta_4 = 49.5$.

7.4 Synchronization of Layered Coding Systems

As outlined in section 3.5, layered coding has many advantages over non-layered coding. The encoder and the decoder delays of each layer may be equal but they can incur different delays and jitters as they may be transmitted in different paths. The delays and jitters of the two layered streams can be represented as shown in figure 7.4.

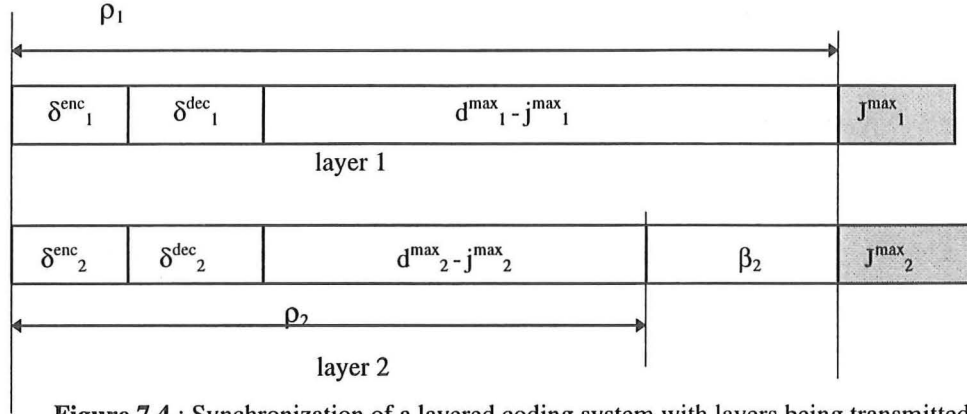


Figure 7.4 : Synchronization of a layered coding system with layers being transmitted across different paths

In synchronizing layered coding systems, the allowable skew tolerance between the two streams is zero. After adding the static delay, as both layers are produced and presented to the user by the same sender clock and receiver clock respectively, the arrival and display of frames can be depicted as shown in figure 7.5.

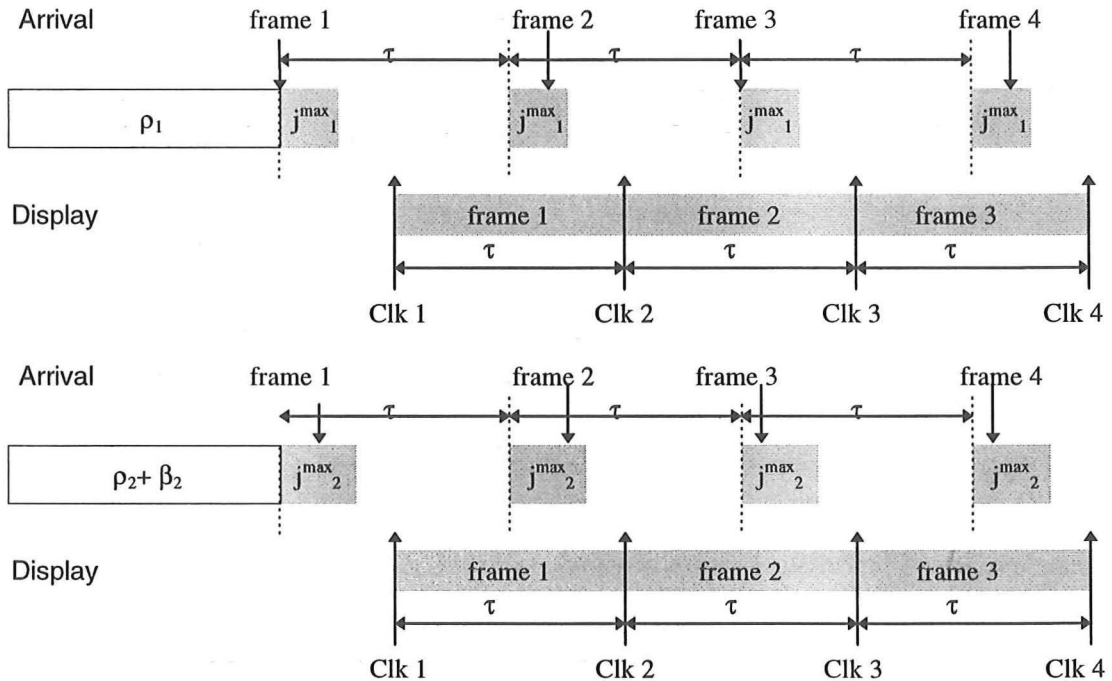


Figure 7.5 : Arrival times and display times of frames

Although, the layered streams may be synchronized most of the time, one layer may slip a frame while the other layer may not, as shown in figure 7.6. In this situation, the synchronization between the layers is offset by one frame. This is called *single frame asynchrony* between the layers. If the second layer slips a frame at the beginning of its uncertainty region (section 4.5) while the first layer only slips a frame at the end the uncertainty region, the frame asynchrony will remain until the receiver clock drifts by j_1^{\max} . Thus, by applying the SDC mechanism alone, uninterrupted synchronization between the layered streams cannot be achieved, and the synchronization mechanism should be extended to ensure synchronization of layered streams.

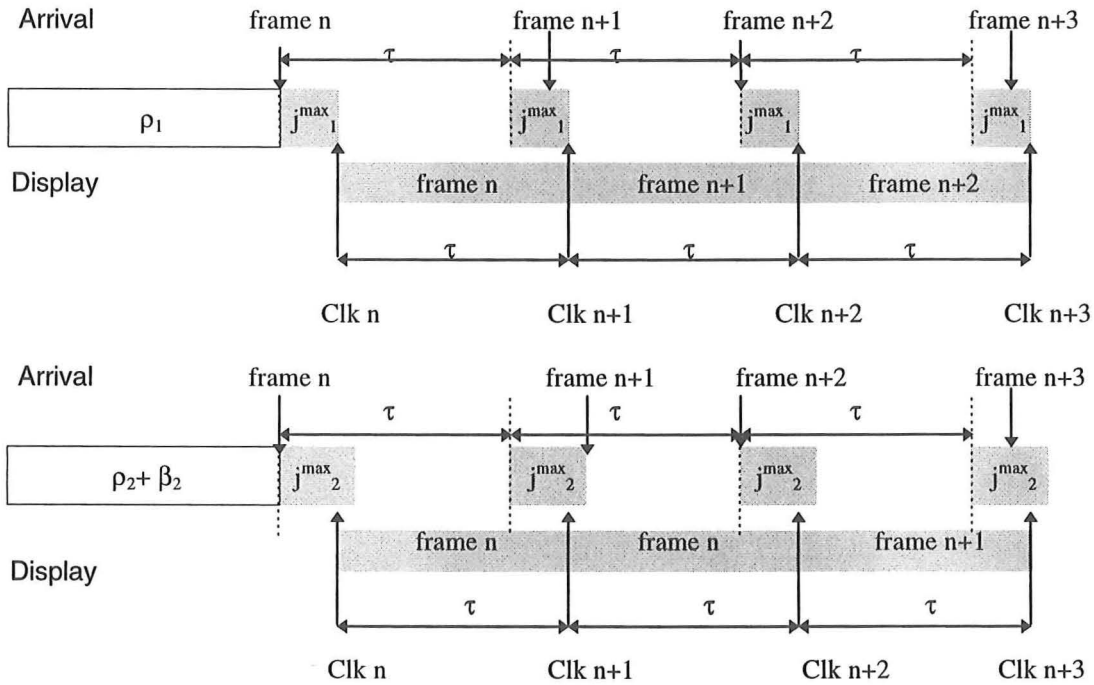


Figure 7.6 : Frame asynchrony between layer 1 and layer 2

Typically, for a sender to receiver clock frequency deviation of 100 ppm, and the j^{\max} value is equivalent to 500 μ s, the clock would drift by j^{\max} in about 5 seconds. Thus, the time interval of single frame asynchrony is about 5 seconds, in the above situation. A media frame will be slipped or paused once every 10000 frames or 333 s, if the sender to receiver clock asynchrony is equal to 100 ppm (section 4.6). Thus, a single frame asynchrony between the layers would exist only 1.5% ($5/333 \times 100$) of the time. Therefore, 98.5% of the time, there is no frame asynchrony between the two layers. The above proportion remains the same for a different sender to receiver clock frequency deviation but is very dependent on the transfer delay-jitter (j^{\max}).

In order to eliminate the effect due to single frame asynchrony, two approaches are proposed.

1. Ignoring the higher layer coded information during the period of frame asynchrony.
2. Sending a control message between the base layer and the higher layer when one layer either slips or pauses a frame, to inform the other layer to follow suit.

In the first approach, the information loss will not be considerable, since firstly the time period is small and secondly the base layer is still decoded and presented. The second approach, is still superior than most of the proposed inter-media synchronization mechanisms, as time stamps or control messages are not sent across the network. The mechanism proposed in the second approach, exchanges no control information between the layers typically 98.5 % of the time and a message is exchanged between the layers, only when a slip or a pause occurs. However, other checks are necessary to ensure that in the other layer, a slip does not occur after a pause or vice versa.

7.5 Example of Synchronizing a Layered coding System

Assume two layers of a layered coding system are transmitted as shown in figure 7.7. Layer 1 would have a delay-jitter (j^{\max}_1) of 1 ms as it is transmitted across four ATM stages. Similarly, layer 2 would have a delay-jitter (j^{\max}_2) of 0.75 ms. The delays and jitters of the two layers are represented in figure 7.8.

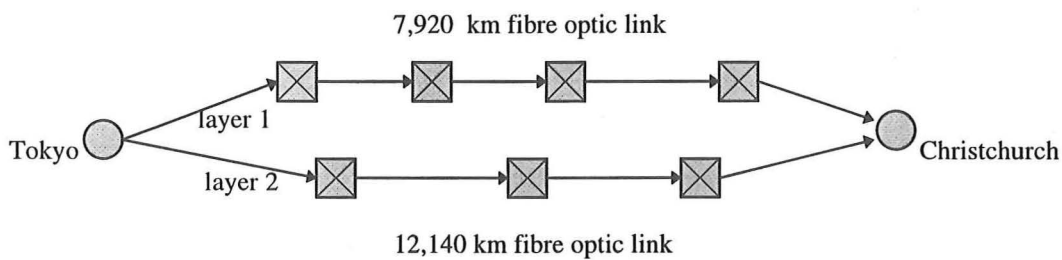


Figure 7.7 : Synchronizing a layered coding system

The two layers can be synchronized by adding a static delay of 21.1 ms to stream 1. If the sender to receiver clock asynchrony is 100 ppm then the uncertainty intervals of layer 1 and 2 are equivalent to 10 s and 7.5 s, respectively. Therefore, maximum frame asynchrony occurs over a period of 10 s every 333 s. Therefore a single frame asynchrony will exist

3% of the time. During the period of single frame asynchrony, one of the proposed approaches may be used to ensure synchronization.

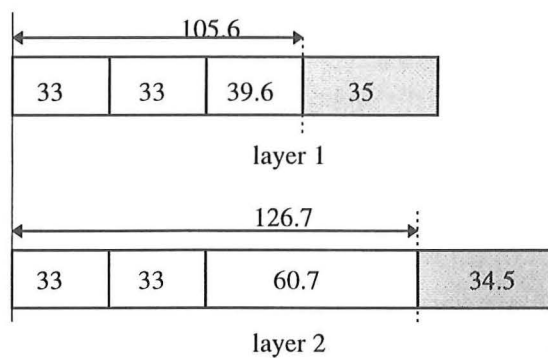


Figure 7.8 : End-to-end delays and jitters of the two layers

CHAPTER 8

DISCUSSION AND CONCLUSION

This chapter presents a discussion of the achievements in this thesis and a conclusion. Further work that could be carried out in this field is also discussed.

8.1 Discussion

In this thesis, a synchronization technique that could be used to guarantee QoS requirements in B-ISDN is illustrated. Almost all of the aspects of synchronization QoS, are analyzed.

The mechanisms that are developed in the thesis provide a robust framework for guaranteeing synchronization QoS within acceptable tolerances.

Quality of Service as perceived by humans is subjective. The subjective Quality of Service parameters have to be converted into objective values to specify, evaluate and maintain them. Guaranteed QoS provision is one of the main features of Broadband ISDN. In providing guaranteed Quality of Service to the users of B-ISDN, a comprehensive Quality of Service management framework should be in place. The QoS management has to integrate the processes of QoS negotiation and translation, QoS verification, QoS maintenance, resource management and signalling.

A structured layered model is used to incorporate all the processes of QoS management. In QoS negotiation, a service contract is established between the user and the service provider. The agreed terms of the service contract may also be called user QoS. The user QoS is not directly perceivable by the network and need to be translated to network QoS. The QoS verification function monitors the performance at various layers while QoS maintenance function identifies QoS violations and tries to rectify them. Resource management does the management of all the resources which includes bandwidth and special resources. Signalling is the message transfer syntax used in QoS management. In the QoS management framework, establishing and guaranteeing synchronization of bearer connections is the responsibility of special resource management, which is a sub-process of resource

management. The special resources that need to be configured and their synchronization characteristics are discussed in detail.

A mechanism to ensure intra-media synchronization is proposed in this thesis, taking into account the jitter in the end-to-end delay as well as the lack of synchronization between sender and receiver clocks. A slip buffer is used to counteract the above mentioned effects. If a media frame arrives too late to be decoded at the slip buffer, the previous frame is paused, whereas if a media frame arrives too early, i.e. the slip delay exceeds a particular threshold, a media frame is slipped. The slip/pause method ensures intra-media synchronization. The intra-media synchronization mechanism is used to develop robust a priori estimates for the minimum and maximum bounds of delay. This robust delay behaviour can be used in facilitating inter-media synchronization.

Ideally, two streams produced by a single source such as audio and video should be perfectly synchronized during playout. However, it is impossible to achieve perfect synchronization between two media streams as there is an uncertainty or jitter in any media stream. Furthermore, in synchronization between audio and video, video can lead /lag audio within skew tolerances, without noticeable quality degradation. Inter-media synchronization is the process of synchronizing media streams within acceptable skew tolerances.

Broadband ISDN utilizes ATM as the underlying transfer mode and fibre optics as the physical transfer medium. When the above underlying technologies are used the end-to-end delay-jitter is low and it is normally possible to synchronize audio and video. Thus effective audio and video inter-media synchronization QoS guarantees could be provided for multimedia calls between any two parts of the world.

In this thesis, the mechanism of Static Delay Compensation is proposed to ensure inter-media synchronization between two or more media streams. In synchronizing two media streams, two skew constraints need to be considered. If the media streams satisfy both constraints then the streams are synchronized already. If one skew constraint is satisfied the streams may be synchronized by adding a static delay to one of the streams. It was shown in the examples that if fibre optics are used as the physical transfer medium, audio / video could be synchronized for the above two cases. However, given a certain relationship between skew constraints, it may be impossible to synchronize the streams, with only one skew constraint satisfied. If both skew constraints are not satisfied the media streams cannot

be synchronized. Although probability of this situation occurring is very low if fibre optics were used for transmission, it was also shown in an example that if satellite communications were used for transmission, audio and video are generally not able to be synchronized.

In synchronizing a layered coding system, SDC mechanism should be applied by adding static delays to the layers with lower end-to-end delay. Although by simple application of SDC, synchronization between the layered streams can be achieved most of the time, there are short time intervals where a single frame asynchrony may exist between the layers.

In an application where multiple audio and video streams are transmitted all media may have to be played out in synchrony. Also in other specialized applications either multiple streams originating from different sources may have to be played in synchrony at a single destination or a single stream may have to be played in synchrony in multiple destinations. The mechanism developed to synchronize two media streams is extended to synchronize multiple media streams. In synchronizing multiple media streams there is a number of other constraints that need to be satisfied and these constraints are discussed in detail. An equation is derived to calculate the static delays to be added to any media stream when synchronizing three media streams. This equation is extended to calculate the static delays to be added to media streams when synchronizing n media streams.

SDC is a simple and a flexible mechanism that could be utilized to ensure inter-media synchronization, as media streams need neither to be multiplexed together nor routed through a common path. It has the capability of guaranteeing synchronization QoS while optimizing resource utilization. It reduces the use of bandwidth by being able to accommodate optimized routing topologies, while it reduces the usage of playout buffer by selecting the lowest static delay. Also SDC is a mechanism that exploits the elegant characteristics of the underlying transfer mechanisms of B-ISDN. In SDC, once the media streams are set up by adding the necessary static delays there are no exchanges of control information between either the sender and the receiver or the architectural layers of the QoS management framework.

8.2 Conclusion

In this thesis, a novel scheme called Static Delay Compensation (SDC) was proposed for

inter-media synchronization, based on skew tolerances. SDC is a simple and a flexible mechanism that guarantees synchronization QoS while optimizing the usage of resources. It reduces the protocol overhead by eliminating the transfer of control messages once the connections are setup. If fibre optics and ATM switches are used as the underlying technologies, inter-media synchronization can be generally guaranteed by applying the mechanism of Static Delay Compensation.

Derivation of constraints to be satisfied when synchronizing any number of media streams was presented. Generalized equations were derived to calculate the static delays that should be added to media streams when synchronizing any number of media streams. Static Delay Compensation is a very comprehensive mechanism that could be used to guarantee synchronization QoS in broadband communications.

8.3 Future Considerations

This thesis primarily presents a theoretical basis for guaranteeing synchronization QoS. In a practical implementation, although the basis outlined in this thesis may still remain valid, there may be additional factors that need to be satisfied. Static Delay Compensation should be tested in a real network to derive these factors. Thus, a practical analysis is necessary to extend the work carried out in this thesis. A technique such as protocol synthesis may be used to translate the high-level QoS management framework into low-level hardware and software implementations.

SDC should be extended for the synchronization of a layered coding system. In synchronizing layered coding systems, two approaches were proposed to circumvent the possibility of single frame asynchrony. These were; to ignore the higher layer coded information or sending control messages between the layers. Experiments should be carried out to derive the information loss of the first approach and the protocol overhead involved in the second approach.

The SDC mechanism may use the viewing distances as a variable in setting up multimedia calls. The relationship of allowable skew with distance should be found experimentally and the results can be incorporated into the SDC mechanism. The SDC mechanism may also be tested with other specialized environments such as tele-surgery, remote trouble shooting or remote aircraft maintenance to derive any extensions that are required.

REFERENCES

AHMED H. (1989), "Adaptive Terminal Synchronization in Packet Data Networks", Proc. *IEEE GLOBECOM'89*, pp 728- 731

ANDERSON D. AND HOMSEY G. (1991), " A Continuous Media I/O Server and Its Synchronization Mechanism", *IEEE Computer*, October, pp 51-57

BAUER D., HUMER-HAGER T., PUTZ K. (1995), "Field Trials of Multimedia Applications in ATM Networks", *Telecom Report International 18*, April, pp 24-28.

BIGGAR M. AND JOHNSON A. (1990), "The Challenge of the Universal Video Codec", Proc. *Australian Video Communications Workshop*, July, pp 175-183

CAMPBELL A., COULSON G., GARCIA F. AND HUTCHISON D. (1992), " A Continuous Media Transport and Orchestration Service", Proc. *ACM SIGCOMM'92*, Maryland, USA

CAMPBELL A., COULSON G., GARCIA F. AND HUTCHISON D.(1993a), " Resource Management in Multimedia Communication Stacks", Proc. *4th IEE Conference on Telecommunications*, Manchester UK, April, pp 1-9.

CAMPBELL A., COULSON G., GARCIA F., HUTCHISON D. AND LEOPOLD H.(1993b), "Integrated Quality of Service for Multimedia Communications", Proc. *IEEE INFOCOM'93*, pp 732-739.

CAMPBELL A., COULSON G. AND HUTCHISON D. (1994), "A Quality of Service Architecture", Proc. *ACM SIGCOMM '94*, pp 6-27.

CCITT Recommendation H.261

CCITT Recommendations G.711, G.721, G.722 and G.728

CHANG S-F AND MESSERSCHMITT D. (1993), "A New Approach to Decoding and Compositing Motion Compensated DCT-based Images", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-93)*, pp 421 - 424

CHAIRIGLIONE L. (1995), "MPEG A Technological Basis for Multimedia Applications", Liu P. (Eds.), *IEEE Multimedia*, Spring, pp 85-89

COHENNEC J., ADAM P. AND HOUDOIN T. (1985), "Asynchronous Time Division Networks : Terminal Synchronization for Video and Sound Signals", *Proc. IEEE GLOBECOM'85*, November, pp 791-794

COOPER J. (1988), "Video-to-audio Synchrony Monitoring and Correction", *Society of Motion Picture and Television Engineers Journal*, August, pp 695-698.

COPPISETTI N., KWATRA S. AND AL-ASMARI A. (1993), "Low-Complexity Subband Encoding for HDTV Images", *IEEE Journal on Selected Areas in Communications*, Vol 11, No 1, January, pp 77-87

DE PRYCKER M., RYCKEBUSH M. AND BARRI P. (1987), "Terminal Synchronization in Asynchronous Networks", *Proc. International Communications Conference*, Seattle, pp 800-807

DE PRYCKER M. (1993), *Asynchronous Transfer Mode: Solution for Broadband ISDN*, Second Edition, Ellis Horwood Ltd., England

DIMOLISTAS S., COCORAN F. AND PHIPPS J. (1993), "Impact on Transmission Delay on ISDN Videotelephony", *Proc. IEEE GLOBECOM'93*

ESCOBAR J., DEUTSCH D. AND PARTRIDGE C. (1992), "Flow Synchronization Protocol", *Proc. IEEE GLOBECOM '92*.

GRAHAM C. (1965), *Vision and Visual Perception*, GRAHAM C. (eds.), John Wiley & Sons Inc., First Edition, Chapter 20, pp 581.

- HANDEL R. AND HUBER M. (1991), *Integrated Broadband Networks : An Introduction to ATM-based networks*, Addison-Wesley Publishing Company
- HOLLIER C. AND ARNOLD J. (1990), " A Hierarchical Sub-band Image Coding Scheme for Videophone Images on a Broadband ISDN", *Australian Video Communications Workshop*, July, pp 185-191
- HU YIHONG (1995), *A Transport Protocol for Multimedia Multicast Communications*, Master's Thesis, University of Canterbury, May.
- HUTCHISON D., COULSON G., CAMPBELL A. AND BLAIR G. (1994), "Quality of Service Management in Distributed systems", *Distributed Systems Management*, SLOMAN M.(eds), Imperial College London, Chapter 11, pp 1-31.
- IAI S., KURITA T. AND KITAWAKI N. (1993), "Quality Requirements Multimedia Communication Services and Terminals - Interaction of Speech and Video Delays", Proc. *IEEE GLOBECOM'93*, pp 394 -398
- JEFFAY K., STONE D. AND SMITH F. (1994), "Transport and Display Mechanisms for Multimedia Conferencing Across Packet Switched Networks", *Computer Networks and ISDN Systems*, Vol. 26, No. 10, pp 1281-1304
- JUNG J. AND SERET D. (1993), "Translation of QoS Parameters into ATM Performance Parameters in B-ISDN", Proc. *IEEE INFOCOM'93*, pp 748-755
- KATSEFF H. AND GAGLIANELLO R. (1991), " On the Synchronization and Display of Multiple Full Motion Video Streams", Proc. *TRICOMM'91*, pp 3-9
- KIM JIN TAE, LEE HYAN JOO AND CHAOI JONG SOO (1993), "Subband Coding Using Human Visual Characteristics for Image Signals", *IEEE Journal on Selected Areas in Communications*, Vol 11, No 1, January, pp 59-63
- LAU R. AND FLEISCHER P. (1992), "Synchronous Techniques for Timing Recovery in B-ISDN", Proc. *IEEE GLOBECOM '92*, pp 814-819

- LEE T.(1988), "Non-Blocking Copy Networks for Multicast Packet Switching", *IEEE Journal on Selected Areas in Communications*, Vol 6, No 9, December, pp 1455-1467
- LEI S-M, CHEN T-C AND SUN M-T (1994), "Video Bridging Based on H.261 Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol 4, No 4, August, pp 425-436
- LIM H. , PANG K. DUNSTAN S. AND NGAN K. (1992), "Interworking Video Services with Layered Coding", Proc.. *TENCON'92*, pp 918-922
- LITTLE T. AND KAO F. (1992), " An Intermedia Skew Control System for Multimedia Data Presentation", Proc. *Third International Workshop on Network and Operating System Support for Digital Audio and Video*, November, pp 121-132
- MOUCHTARIS P., MOYER S., FLINGHBAUGH G., LELAND W., POMPLUN D. AND WEINRIB A.(1994), "An Object Oriented Interface for a Distributed Connection Manager", *IEEE*, pp 273-277
- MOURELATOU K., BOULOUTAS A. AND ANAGNOSTOU M. (1994), "An Approach to Identifying QoS Problems", *Computer Communications*, Vol 17, no 8, August, pp 563-570
- MOYER S. AND ROUSE D. (1993),"A Connection Manager for Flexible Specification and Transport Location of Special Resources", *IEEE GLOBECOM'93*, pp 1526-1530
- NAHRSTEDT K. AND SMITH J. (1995a), "The QoS Broker", *IEEE Multimedia*, Spring, pp 53-67.
- NAHRSTEDT K. AND STEINMETZ R. (1995b), " Resource Management in Networked Multimedia Systems ", *IEEE Computer*, May, pp 52-63.
- NAISH G. (1994a), "Protocols for Real Time Multimedia Synchronization and Clock Synchronization", *Technical Report*, University of Canterbury, New Zealand, January.
- NAISH G. (1994b), "A Taxonomy and Quality of Service Architecture for Continuous Media Synchronization", *Technical Report*, University of Canterbury, New Zealand, July.

- NAISH. G., WEERASIRI K. AND CAREY-SMITH C. (1995), "Static Delay Compensation; A Mechanism for Synchronizing Two Media Streams in Broadband Communications", *Technical Report*, University of Canterbury, New Zealand, August.
- PAGLIALUNGA A., RONCHETTI L., FAGLIA L. AND MINZER S. (1994), "The Control of Network Resources for Complex B-ISDN Services", *Second International IFIP Conference in Broadband Communications*, Paris-France, March, pp 1-15
- PANCHA P. AND EL ZARKI M. (1994), "MPEG Coding for Variable Bit Rate Video Transmission", *IEEE Communications Magazine*, May, pp 54-66
- PINGALI S. AND KUROSE J. (1991), "On Scheduling Two Classes of Real-Time Traffic with Identical Deadlines", *Proc. IEEE GLOBECOM'91*, pp 460-465
- RAVINDRAN K. AND BANSAL V. (1993), "Delay Compensation Protocols for Synchronization of Multimedia Data Streams", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 5, No. 4, August, pp 574-589
- RESEARCH FOR ADVANCED COMMUNICATIONS IN EUROPE (RACE) Project QOSMIC - deliverable D 1.3 C (1992), "QoS and Performance Relationships", QOSMIC Consortium, June.
- RESEARCH FOR ADVANCED COMMUNICATIONS IN EUROPE (RACE) Project MAGIC- deliverable 10 A (1994), "Protocols and Concepts of B-ISDN signalling", September.
- RESEARCH FOR ADVANCED COMMUNICATIONS IN EUROPE (RACE) Project MAGIC- deliverable 14 (1995), "B-ISDN signalling Performance : Issues and Analysis", September.
- SHACHAM N. (1992), "Multipoint Communication by Hierarchically Encoded Data", *Proc. IEEE INFOCOM'92*, PP 2107-2114
- SHAE Z-Y AND CHEN M-S (1993), "Mixing and Playback of JPEG Compressed Packet Videos", *Proc. IEEE GLOBECOM'92*, pp 245-249

SINGH A. AND BOVE V. (1993), "Multidimensional Quantizers for Scalable Video Compression", *IEEE Journal on Selected Areas in Communications*, Vol 11, No 1, January, pp 36-43

SINGH R., LEE S. AND KIM C. (1988), "Jitter and Clock Recovery for Periodic Traffic in Broadband Packet Networks", Proc. *IEEE GLOBECOM'88*, pp 1468-1473

SRIRAM K., MCKINNEY R. AND SHERIF M. (1991), "Voice Packetization and Compression in Broadband ATM Networks", *IEEE Journal on Selected Areas in Communications*, Vol 9 No 3, April, pp 294-304

STEINMETZ R. AND ENGLER C. (1993), "Human Perception of Media Synchronization", *Technical Report 43:9310*, IBM European Networking Centre, November

VENKAT RANGAN P., VIN H. AND RAMANATHAN S. (1993), "Communication Architectures and Algorithms for Media Mixing in Multimedia Conferences", *IEEE /ACM Transactions on Networking*, Vol 1, No 1, pp 20-30

WILLEBEEK-LEMAIR M., KANDLUR D. AND SHAE Z -Y (1994), "On Multipoint Control Units for Videoconferencing", Proc. *19th conference on Local Computer Networks*, pp 356-364

ZAFAR S., ZHANG Y-Q AND JABBARI B. (1993), "Multiscale Video Representation Using Multiresolution Motion Compensation and Wavelet Decomposition", *IEEE Journal on Selected Areas in Communications*, Vol 11, No 1, January, pp 24-35

ZEBARTH J. (1993), "Let Me Be Me", Proc. *IEEE GLOBECOM'93*, pp 389-393